



Geodesic flow of the averaged controlled Kepler equation

Bernard Bonnard, Jean-Baptiste Caillau

► To cite this version:

Bernard Bonnard, Jean-Baptiste Caillau. Geodesic flow of the averaged controlled Kepler equation. 2007. hal-00134702v1

HAL Id: hal-00134702

<https://hal.science/hal-00134702v1>

Preprint submitted on 5 Mar 2007 (v1), last revised 22 Feb 2008 (v6)

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.

GEODESIC FLOW OF THE AVERAGED CONTROLLED KEPLER EQUATION*

BERNARD BONNARD[†] AND JEAN-BAPTISTE CAILLAU[‡]

Abstract. A normal form of the Riemannian metric arising when averaging the coplanar controlled Kepler equation is given. This metric is parameterized by two scalar invariants which encode its main properties. The restriction of the metric to \mathbf{S}^2 is shown to be conformal to the flat metric on an oblate ellipsoid of revolution, and the associated conjugate locus is observed to be a deformation of the standard astroid. Though not complete because of a singularity at infinity in the space of ellipses, the metric has convexity properties that are expressed in terms of the aforementioned invariants, and related to surjectivity of the exponential mapping. Optimality properties of geodesics of the averaged controlled Kepler system are finally obtained thanks to the computation of the cut locus of the restriction to the sphere.

Key words. orbit transfer, Riemannian metrics, cut and conjugate loci

AMS subject classifications. 49K15, 70Q05

Introduction. Modern orbit transfer missions, *e.g.* interplanetary ones, use electro-ionic propulsion [8], that is engines with very low thrust compared to chemical boosters. Typical performance indexes for these transfers include time [7] or fuel minimization [12, 17]. The underlying model for such missions is the standard two-body system (swingby or planetary capture phenomena being also treated in this framework, for instance using several phases [2]) described by the controlled Kepler equation. In its simplest form, it is a second order control system involving the position vector q in \mathbf{R}^3 , and the thrust of the engine, u ,

$$\ddot{q} = -q/|q|^3 + u. \quad (0.1)$$

More realistic models take into account the variation of the mass due to fuel consumption, higher order terms of the potential, *etc.*

Let $c = q \times \dot{q}$ denote as usual the angular momentum, $E = \dot{q}^2/2 - 1/|q|$ the mechanical energy, and define the *elliptic domain*, Q , according to

$$Q = \{(q, \dot{q}) \mid c \neq 0, E < 0\}.$$

On this domain, the unperturbed motion of (0.1) is known to be an ellipse, and the standard first integrals of the movement can be used as geometric coordinates. In the coplanar case where the control remains into the plane defined by the initial angular momentum, the geometry of the osculating ellipse is determined by three such first integrals, for instance the semi-major axis a —which can be replaced by the mean movement, $n = a^{-3/2}$ —, the eccentricity, e , and the argument of perigee, ω (see [15]). These coordinates reveal the structure of the coplanar elliptic domain which is fibered over \mathbf{S}^1 since $Q = X_e \times \mathbf{S}^1$, with X_e the *space of ellipses*,

$$X_e = \{(n, e, \omega) \mid n > 0, 0 < e < 1, \omega \in \mathbf{S}^1\} = \mathbf{R}_+^* \times \mathbf{D},$$

*This work was done in the framework of the HYCON Network of Excellence, contract number FP6-IST-511368.

[†]Institut de Mathématiques de Bourgogne, UMR CNRS 5584, 9 avenue Savary, F-21078 Dijon (bernard.bonnard@u-bourgogne.fr).

[‡]ENSEEIH-IRIT, UMR CNRS 5505, 2 rue Camichel, F-31071 Toulouse (caillau@n7.fr).

where \mathbf{D} is the open unit disc in the complex plane. One has then to add to the previous three first integrals a coordinate defining the position on the osculating ellipse, namely the longitude, l in \mathbf{S}^1 .

Numerical experiments in the minimum time case [5] give evidence for the existence of conjugate points where extremals of the system cease to be locally minimizing [20]. A simplified approximation of the system is then desirable to study optimality properties. For low thrust transfers, the modulus of the control is very small and many revolutions are required to achieve an orbit transfer, *e.g.* from a low eccentric initial orbit towards the geosynchronous one [7]. As a consequence, while the geometric coordinates arising as first integrals of the unperturbed motion are slowly varied, the longitude acts as a fast variable, and the small parameter of the problem is the inverse of the final value of the longitude cumulated along the whole trajectory (angular length of the transfer). Averaging with respect to longitude has been considered in [11], and it is shown in [4] that the averaged Hamiltonian of the minimization of the energy—the L^2 -norm of the control—remains integrable (see also §2). More precisely, this averaged Hamiltonian is quadratic in the adjoint state and associated to a Riemannian metric in orthogonal form and with singularities on X_e ,

$$g = dn^2/(9n^{1/3}) + 2n^{5/3}de^2/(5(1-e^2)) + 2n^{5/3}e^2d\omega^2/(5-4e^2). \quad (0.2)$$

The singularities are $n = 0$, which is a point at infinity in the space of ellipses (*i.e.* $a = \infty$), and $e = 1$, which defines the parabolic boundary of the domain, while $e = 0$ is only due to the use of polar coordinates on \mathbf{D} and removed by taking cartesian ones instead. The effect of the singularity at infinity is discussed in §3.

The first section of the paper is devoted to curvature computations. After a partial compactification, a normal form of the metric is obtained, revealing two scalar invariants. The first one is related to convexity issues, whereas the second defines the geometry of the restriction of the metric to \mathbf{S}^2 , measuring in particular the number of closed simple geodesics on the sphere (see §2). We prove in the second section that the metric is integrable in the class of harmonic functions, and provide explicit quadratures in suited coordinates. An estimation of the length of closed geodesics combined with the curvature evaluation of §1 allows to compute the injectivity radius of the metric restricted to \mathbf{S}^2 and to devise a necessary condition for optimality of metrics in the normal form derived. In Kepler's case, this condition is not fulfilled and cut points exist for the partially compactified metric. Section 3 deals with convexity issues. Though the singularity at infinity of the metric results in incompleteness, convexity may occur and is characterized thanks to the two aforementioned parameters. The analysis is reduced to a discussion in two-dimensional meridian half-planes, and related to surjectivity of the exponential mapping. In the last section we compute the cut locus of the metric restricted to the sphere, and the astroid-like associated conjugate locus is obtained numerically [6]. Both are related to the separating line and conjugate locus of the full three-dimensional metric.

1. Curvature of the system. We begin with a partial compactification of the space of ellipses, X_e , which absorbs the singularity at the parabolic boundary, $e = 1$.

Let $r = (2/5)n^{5/6}$, and let (θ, φ) be the usual angular coordinates on the two-sphere (*i.e.* we set $e = \sin \varphi$, $\theta = \omega$, and lift \mathbf{D} to \mathbf{S}^2). We get the following.

PROPOSITION 1.1. *A normal form of Kepler's metric (0.2) is*

$$g = dr^2 + (r^2/c^2)(G(\varphi)d\theta^2 + d\varphi^2)$$

which is defined on $X = \mathbf{R}_+^* \times \mathbf{S}^2$ with

$$G(\varphi) = \sin^2 \varphi / (1 - (1 - \mu^2) \sin^2 \varphi)$$

and positive $c, \mu, \mu \leq 1$. In Kepler's case, $c = \sqrt{2/5}$ and $\mu = 1/\sqrt{5}$.

Again, the singularities $\varphi = 0$ (π) at the poles are simply due to the choice of coordinates on \mathbf{S}^2 , and $r = 0$ is the only singularity left since $e = 1$ is absorbed as the equator, $\varphi = \pi/2$.

The restriction of the metric to $\{r = c\} \simeq \mathbf{S}^2$ is $g_2 = G(\varphi)d\theta^2 + d\varphi^2$, and we denote similarly the resulting normal forms of Hamiltonians:

$$H = p_r^2/2 + (c^2/r^2)H_2, \quad H_2 = (1/2)(p_\theta^2/G(\varphi) + p_\varphi^2). \quad (1.1)$$

An analytic metric on the two-sphere of revolution like g_2 is well known to be isometric to $f(z)g_0$, where $g_0 = \sin^2 \varphi d\theta^2 + d\varphi^2$ is the restriction of the flat metric to \mathbf{S}^2 and f a positive function of the vertical coordinate z (see [3]). A more specific construction is available here.

PROPOSITION 1.2. *The metric g_2 is conformal to the flat metric restricted to an oblate ellipsoid of revolution of unit semi-major axis and semi-minor axis μ .*

Proof. We have indeed

$$g_2 = G(\varphi)d\theta^2 + d\varphi^2 = g_1/(1 - (1 - \mu^2) \sin^2 \varphi),$$

where $g_1 = \sin^2 \varphi d\theta^2 + (1 - (1 - \mu^2) \sin^2 \varphi)d\varphi^2$ is the restriction of the flat three-dimensional metric to the ellipsoid parameterized by

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

whence the result. \square

Accordingly, there is a natural homotopy from g_2 to the flat metric on \mathbf{S}^2 , having the parameter μ varying up to $\mu = 1$ (since $g_2|_{\mu=1} = g_1|_{\mu=1} = g_0$). This resemblance with the ellipsoid of revolution defined by μ is crucial to understand the cut and conjugate loci computations of §4.

When $r \rightarrow \infty$, the semi-major axis coordinate $a = (5r/2)^{-4/5}$ tends to zero, the collision point in the space of ellipses. We analyse the effect of the collision on the curvature.

Let V be the span of $\partial/\partial\theta$ and $\partial/\partial\varphi$. The sectional curvature of K_V of g is $R_{2323}/|\partial/\partial\theta \wedge \partial/\partial\varphi|^2$ where R is the Riemannian curvature tensor and $R_{2323} = R(\partial/\partial\theta, \partial/\partial\varphi, \partial/\partial\theta, \partial/\partial\varphi)$. Quite clearly, g is flat if and only if K_V is zero.

PROPOSITION 1.3. *The sectional curvature of g in Kepler's case is*

$$K_V = (1 - 24 \cos^2 \varphi - 16 \cos^4 \varphi)/(r^2(1 + 4 \cos^2 \varphi)^2) \rightarrow 0, \quad r \rightarrow \infty,$$

and the metric is asymptotically flattened by the collision.

We show in §4 that the analysis of optimality of the geodesic flow of g is deduced from the properties of the restriction to \mathbf{S}^2 . We immediately compute the Gauss curvature of g_2 , $K = -(d^2\sqrt{G}/d\varphi^2)/\sqrt{G}$.

PROPOSITION 1.4. *The Gauss curvature of g_2 is*

$$K = (\mu^2 - 2(1 - \mu^2) \cos^2 \varphi)/(1 - (1 - \mu^2) \sin^2 \varphi)^2.$$

The curvature reaches its maximum, $K = 1/\mu^2$, along the equator.

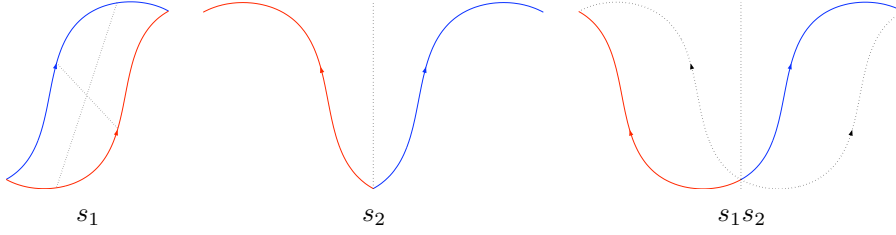


FIG. 2.1. Action of the Klein group on geodesics.

By Rauch theorem, the first conjugate time along any geodesic on \mathbf{S}^2 is then bounded below by $\mu\pi$, and this bound is optimal since it is reached on the equator $\varphi = \pi/2$ where the curvature is constant. In particular, the distance from any point on the sphere to its cut locus is hence bounded over by $\mu\pi$, and the injectivity radius, $i(\mathbf{S}^2)$, which is the infimum of such distances over the whole manifold, shares the same bound: $i(\mathbf{S}^2) \leq \mu\pi$. Moreover, it is known [10] that this radius is reached—the manifold is compact—either at a conjugate point, or at the half of a simple closed geodesic. The task of computing the closed geodesics on the sphere is completed in next section, thus providing a first necessary condition for global optimality of g thanks to the estimation of the injectivity radius.

2. Integrability. The unperturbed Kepler motion is classically integrable, and so remains the averaged controlled one, independently of the choice of parameters c , μ in (1.1).

PROPOSITION 2.1. *The coordinate θ is cyclic, and H , H_2 , p_θ are three independent first integrals in involution. On \mathbf{S}^2 , the linear first integral p_θ verifies the Clairaut relation [3, 9], $p_\theta = \cos(\phi)\sqrt{G(\varphi)}$, where ϕ is the angle of the geodesic with a parallel.*

The geodesic flow on $X = \mathbf{R}_+^* \times \mathbf{S}^2$ is thus Liouville-integrable which we can also check by a direct computation of r on the level set $H = 1/2$ (parameterization of geodesics by arc length).

LEMMA 2.2. *The coordinate r^2 is a degree two polynomial depending only on r_0 and p_{r_0} , $r^2 = t^2 + 2r_0p_{r_0}t + r_0^2$.*

The integration is then performed using the time change $d\tau = dt/c^2$, and one readily gets

$$\tau(t, r_0, p_{r_0}) = c^2(\arctan(t/(r_0 \cos \alpha_0) + \tan \alpha_0) - \alpha_0)/(r_0 \cos \alpha_0) \quad (2.1)$$

with $p_{r_0} = \sin \alpha_0$. Whenever $p_{r_0} = \pm 1$, the angles θ and φ are constant on $H = 1/2$, and we set $\tau = 0$. Parameterizing again extremals by arc length on $\{r = c\} \simeq \mathbf{S}^2$, we set $H_2 = 1/2$ and proceed to the integration of g_2 , first underlying the symmetries of the system.

Since θ is cyclic, we can normalize θ_0 to zero, and the action of the two reflections $s_1 : p_\theta \mapsto -p_\theta$, $s_2 : p_\varphi \mapsto -p_\varphi$ on the Hamiltonian H_2 is clear: s_1 defines an axial symmetry in the (θ, φ) -plane with respect to $(O\varphi)$, while s_2 is a central symmetry $(\theta, \varphi) \mapsto (2\theta_c - \theta, \pi - \varphi)$, θ_c depending on the geodesic. These two reflections generate the Klein group, $\mathbf{V} \simeq \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, which defines an abelian discrete group of symmetries of H_2 (compare [19]), see Fig. 2.1.

PROPOSITION 2.3. *Geodesics on $H_2 = 1/2$ are parameterized by p_θ , and the two coordinates θ , φ are T -periodic with $T = 4\pi/a$ and $a = 2\sqrt{1 + (1 - \mu^2)p_\theta^2}$. On a*

quarter of period $[t_1, t_1 + T/4]$, for p_{φ_0} nonnegative,

$$\theta = \text{sign}(p_\theta) \left[\arctan(\tan(a(t - t_1)/2)/\sqrt{b}) \right]_0^t - (1 - \mu^2)p_\theta t, \quad (2.2)$$

$$\varphi = \arcsin \sqrt{(1 + b)/2 - (1 - b) \cos(a(t - t_1))/2}, \quad (2.3)$$

with $t_1 = -\pi/(2a) - \arcsin((2 \sin^2 \varphi_0 - 1 - b)/(1 - b))/a$ and $b = p_\theta^2/(1 + (1 - \mu^2)p_\theta^2)$. The quadratures are extended by analyticity on the whole period, and obtained for negative p_{φ_0} using a $-2t_1$ time translation on φ and $\dot{\theta}$,

$$\varphi_{|p_{\varphi_0} < 0}(t) = \varphi_{|-p_{\varphi_0}}(t + 2t_1), \quad \theta_{|p_{\varphi_0} < 0}(t) = \theta_{|-p_{\varphi_0}}(t + 2t_1) - \theta_{|-p_{\varphi_0}}(2t_1).$$

Remark. Inner symmetries on each geodesic imply that, for p_{φ_0} nonnegative, $\varphi(t) = \pi - \varphi(2(t_1 + T/4) - t)$ on $[t_1 + T/4, t_1 + T/2]$, and $\varphi(t) = \varphi(2(t_1 + T/2) - t)$ on $[t_1 + T/2, t_1 + T]$. Similar relations hold for θ , since $\dot{\theta} = p_\theta/G(\varphi)$.

COROLLARY 2.4. *The metric g is integrable in the class of harmonic functions.*

When $p_{\varphi_0} = 0$, we get the equator if $\varphi_0 = \pi/2$, the so-called *pseudo-equators* otherwise.

COROLLARY 2.5. *All geodesics with the exception of meridians, $\theta = \text{cst}$, are pseudo-equators.*

Proof. Excluding meridians, on $H_2 = 1/2$ one always has $0 < p_\theta^2 \leq G(\varphi_0) \leq 1/\mu^2$, and there is φ'_0 in $[0, \pi/2]$ such that $p_\theta^2 = G(\varphi'_0)$. The geodesic is a pseudo-equator for the new initial condition φ'_0 . \square

According to Proposition 2.3, $\dot{\theta}$ is periodic, and the variation of θ which is non-decreasing over a period (we restrict ourselves to $p_\theta \geq 0$ by symmetry) is

$$\Delta\theta = 2\pi(1 - (1 - \mu^2)e_0)$$

for a pseudo-equator of initial condition $e_0 = \sin \varphi_0 > 0$. This expression is also valid for meridians passing through the solvable singularity $e_0 = 0$. Indeed, there are jumps in θ at the poles generating the whole family of meridians $\theta = \text{cst}$ starting from θ_0 normalized to zero: $\theta(\varphi = 0+) = \text{cst}$, $\theta(\varphi = \pi) = \text{cst} + \pi$, $\theta(\varphi = 2\pi) = 2\pi$, for a total variation $\Delta\theta = 2\pi$. As $\Delta\theta/(2\pi) \leq 1$, closed geodesics—including meridians—are at least one period long. For pseudo-equators of initial condition e_0 in $]0, 1]$ or meridians ($e_0 = 0$), $T = 2\pi\sqrt{1 - (1 - \mu^2)e_0^2}$ by Proposition 2.3, so that the half-length of any closed geodesic is bounded below by $\mu\pi$.

THEOREM 2.6. *The injectivity radius of the metric g restricted to $\{r = c\} \simeq \mathbf{S}^2$ is $i(\mathbf{S}^2) = \mu\pi$. A necessary condition for global optimality of g on $X = \mathbf{R}_+^* \times \mathbf{S}^2$ is $c/\mu \leq 1$.*

Proof. According to the discussion at the end of §1, $i(\mathbf{S}^2) \leq \mu\pi$, whence the equality given the lower bound of the half-length of closed geodesics. Let us show that g is not globally optimal if $c > \mu$. The time σ on the level set $H_2 = 1/2$ is τ (see (2.1)), up to a renormalization by $H_2|_{H=1}$, that is

$$\sigma(t, r_0, p_{r_0}) = c(\arctan(t/(r_0 \cos \alpha_0) + \tan \alpha_0) - \alpha_0). \quad (2.4)$$

Then

$$\sup_{|p_{r_0}| < 1} \sup_{t \geq 0} \sigma(t, r_0, p_{r_0}) = c\pi, \quad (2.5)$$

this bound not being reached on any geodesic. Let now γ be a geodesic of length $\mu\pi$ on \mathbf{S}^2 with one cut point. Since $\mu\pi < c\pi$, this geodesic is the projection on \mathbf{S}^2 of a geodesic on X (see Lemma 4.1 in §4) which cannot be globally minimizing because of the cut point on γ . \square

Remark. For the condition to be also sufficient, we would need the metric on X to be complete so as to use the standard structure result on the cut locus (decomposed into conjugate points and points on the separating line, see §4). This is not the case, as will be stated in §3. The limit case is the flat case, $c = \mu = 1$ and $X \simeq \mathbf{R}^3 - \{0\}$ where global optimality holds.

COROLLARY 2.7. *The metric is not globally optimal on X in Kepler's case ($c/\mu = \sqrt{2} > 1$).*

As θ and φ share the same period, closed geodesics on \mathbf{S}^2 are characterized by the fact that $\Delta\theta/(2\pi)$ be rational (except for the equator, $e_0 = 1$, which is the only geodesic with constant φ). Conversely, a geodesic which is pseudo-equator for e_0 such that the quotient is irrational densely fills the strip $[\varphi_0, \pi - \varphi_0]$ on the sphere. In Kepler's case where μ^2 is rational, there is one closed geodesic passing through every rational e_0 since $\Delta\theta/(2\pi) = 1 - (1 - \mu^2)e_0$. It is known [14] that there exist at least three simple closed geodesics, and there are actually infinitely many of them—the meridians—because of the symmetry of revolution. We thus discuss the existence of simple closed geodesics modulo rotations on θ .

PROPOSITION 2.8. *There are exactly $[1/\mu^2]$ simple closed geodesics modulo rotations around the poles on (\mathbf{S}^2, g_2) .*

Remark. The invariant μ thus measures the number of simple closed geodesics on the sphere. The result degenerates for $\mu = 1$, all great circles being meridians for appropriate axes on the flat two-sphere.

Proof. For closed geodesics, $\Delta\theta/(2\pi)$ is rational,

$$1 - (1 - \mu^2)e_0 = p/q,$$

and simple ones are obtained for $p = 1$. Then $e_0 = (q - p)/(q(1 - \mu^2))|_{p=1} \leq 1$, so $1 \leq q \leq 1/\mu^2$. \square

In Kepler's case, $1/\mu^2 = 5$ and there are five classes of simple closed geodesics (see Fig. 2.2) among which meridians and pseudo-equator for $e_0 = 5/6$ have 2π -rational lengths (2π and 4π , respectively). In fact, it turns out that the existence of closed geodesics with length in $2\pi\mathbf{Q}$ is expressed in terms of a standard Diophantine equation, the Pell equation [16]. When μ^2 is indeed rational, such closed geodesics are obtained finding a rational e_0 such that $T/(2\pi) = \sqrt{1 - (1 - \mu^2)e_0^2}$ belongs to \mathbf{Q} , that is solving the quadratic Diophantine equation

$$a^2 - (1 - \mu^2)b^2 = c^2. \quad (2.6)$$

This equation is a generalized Pell equation, parameterized by c . In Kepler's case, (2.6) takes the form $5a^2 - 4b^2 = 5c^2$ (with $e_0 = b/a$), which is reduced to the standard equation

$$x^2 - Dy^2 = 25c^2, \quad (2.7)$$

where $D = 20$ is not a perfect square, and where $a = x/5$ when 5 divides x . For $c = 4$, (2.7) has obvious solution $(x, y) = (30, 5)$. The general solution (x_n, y_n) is then obtained solving the unitary Pell equation $x^2 - 20y^2 = 1$ whose particular solution

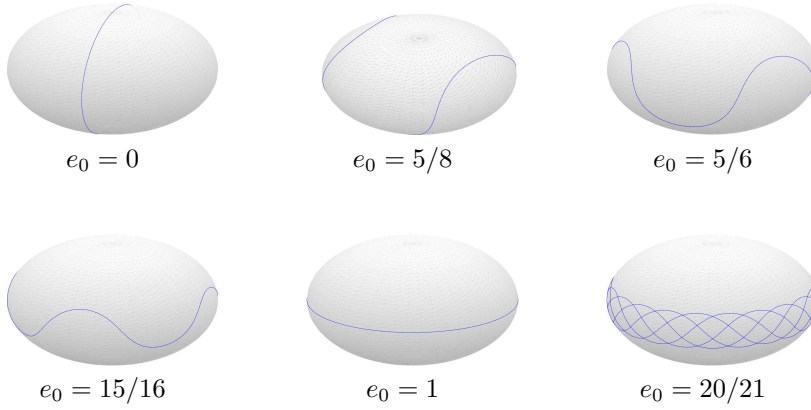


FIG. 2.2. Closed geodesics in Kepler's case. The first five are the simple ones.

$(x, y) = (9, 2)$ acts as a generator: $x_n = 30r_n \pm 100s_n$ and $y_n = 30s_n \pm 5r_n$ with

$$\begin{aligned} r_n &= [(9 + 4\sqrt{5})^n + (9 - 4\sqrt{5})^n]/2, \\ s_n &= [(9 + 4\sqrt{5})^n - (9 - 4\sqrt{5})^n]/(4\sqrt{5}). \end{aligned}$$

To $(x, y) = (30, 5)$ is associated the solution $(a, b) = (6, 5)$ of the original equation (2.6), defining the (simple) closed geodesic passing through $e_0 = 5/6$ of Kepler's case. Infinitely many closed geodesic with 2π -rational length are designed so (*e.g.*, $e_0 = 100/2001$ with $c = 1999$, *etc.*)

Remark. In Kepler's case, meridians are the shorter closed geodesics. Indeed, the length of a closed geodesic such that $\Delta\theta/(2\pi) = p/q$ is $2\pi q\sqrt{1 - (1 - \mu^2)e_0^2} \geq 2\pi\mu q$. If closed geodesics shorter than meridians exist, it is necessarily for $q \leq 1/\mu$. This implies $q = 1$ or 2 in Kepler's case ($1/\mu = \sqrt{5} < 3$), that is $e_0 = 0$ (meridians, precisely) for $q = p = 1$, or $e_0 = 5/8$ for $q = 2, p = 1$, which has length $\pi\sqrt{11} > 2\pi$, since p has to be chosen such that $1 \leq p \leq q, p \wedge q = 1$.

3. Convexity properties. The *meridian half-planes* of X are the subsets $\theta = \text{cst}$. They are all isometric to $X_0 = \{\theta = 0\}$, and the metric g has a flat restriction on them [4]: $g|_{X_0} = dr^2 + (r^2/c^2)d\varphi^2$, that is

$$g|_{X_0} = dr^2 + r^2 d\psi^2, \quad (3.1)$$

setting $\psi = \varphi/c$, φ in $] -\pi/2, \pi/2[$ (upper half-plane). The metric (3.1) is in polar form, and flat coordinates are $x = r \sin \psi$, $z = r \cos \psi$. This reduction is important in Kepler's case since such half-planes correspond to transfers towards circular orbits (where the transversality condition of Pontryagin maximum principle reads $p_\omega = p_\theta = 0$, that is $\theta = \text{cst}$). This is used in practice to initialize, for instance, the computation of transfers towards the geostationary orbit.

The effect of the singularity $r = 0$ is clear in this context.

PROPOSITION 3.1. *The manifold (X, g) is not complete.*

Proof. The separatrices $\varphi = \text{cst}$, that is $\psi = \text{cst}$ in polar coordinates on X_0 , define geodesics on X which reach the singularity $r = 0$ in finite time. \square

Regarding the weaker property of (geodesic) convexity, the following holds.

PROPOSITION 3.2. *A necessary condition for convexity of the manifold (X, g) is $c\mu > 1$.*

Proof. The diameter of (\mathbf{S}^2, g_2) is the half-length of the longest simple closed geodesic which is clearly the equator: $\text{diam}(\mathbf{S}^2) = \pi/\mu$. According to (2.5), the time σ on the sphere is bounded over by $c\pi$. The supremum not being reached, the condition is necessary (no geodesic on X can reach a point that projects into a point on \mathbf{S}^2 further than $c\pi$). \square

Remark. Given two points on X whose projections on the sphere are closer than $c\pi$ from each other, Lemma 4.1 combined with Propositions 4.2, 4.3 of §4 ensures the existence of a geodesic between them without conjugate point or point on the separating line. The sphere is indeed compact, so complete, and one just has to lift a minimizing geodesic onto X thanks to the afore-mentioned lemma. This is not sufficient however to guarantee global optimality of the lifted geodesic since cut points not of the two previous kinds may exist because of incompleteness of the metric on X . It should moreover be noted that the two necessary conditions on global optimality (Theorem 2.6) and convexity are incompatible. In the limit flat case $c = \mu = 1$, $X \simeq \mathbf{R}^3 - \{0\}$ which is clearly not convex.

COROLLARY 3.3. *The metric is not geodesically convex on X in Kepler's case ($c\mu = \sqrt{2}/5 \leq 1$).*

The non-convexity is well depicted in meridian half-planes. In flat coordinates indeed, it is obvious that a geodesic with initial angle ψ_0 cannot reach a point with angle beyond $\pi + \psi_0$ because of the singularity $r = 0$ when $c < 1$ (see Fig. 3.2). Such a pair of points actually projects into points on \mathbf{S}^2 further than $c\pi$, and the analysis in X_0 is sufficient as is now stated.

While optimality properties are related to infinitesimal or global injectivity of the exponential mapping, existence is connected to the presently discussed convexity issues, that is to surjectivity of the exponential. The geodesic flow on the Riemannian manifold X is obtained through the exponential mapping, defined for small times on the whole cotangent space at the initial point, $\exp_{x_0, t} : T_{x_0}^*X \rightarrow X$, by

$$\exp_{x_0, t}(p_0) = \Pi \circ \exp t\overrightarrow{H}(x_0, p_0),$$

where $\exp t\overrightarrow{H}$ is the one-parameter subgroup generated by the Hamiltonian, and $\Pi : T^*X \rightarrow X$ the canonical projection. By homogeneity, we restrict ourselves to the level set $H = 1/2$, parameterizing geodesics by arc length, or consider alternatively $\exp_{x_0} = \exp_{x_0, t=1}$ which, according to Hopf-Rinow, is only defined on an open subset of the cotangent bundle since the manifold is not complete. The same construction holds on the complete manifold \mathbf{S}^2 , and for y_0 on the sphere, we set

$$\Lambda_t = \exp t\overrightarrow{H}_2(\Lambda_0),$$

where $\Lambda_0 = T_{y_0}^*\mathbf{S}^2 \cap \{H_2 = 1/2\} \simeq \mathbf{S}(T_{y_0}^*\mathbf{S}^2)$. Both Λ_0 and Λ_t are Legendrian submanifolds of the fibered space of oriented contact elements [1] which is homeomorphic to the spherical cotangent bundle, $\mathbf{S}(T^*\mathbf{S}^2)$. The *wavefront* at time t is the projection $W(y_0, t) = \Pi(\Lambda_t)$. Its singularities run along the *caustic* of the Lagrangian submanifold $L = \exp_{y_0} t\overrightarrow{H}_2(T_{y_0}^*\mathbf{S}^2)$, that is the set of singular values of Π restricted to L .

According to Proposition 3.2, the surjectivity default of \exp_{x_0} is described by the wavefront $W(y_0, c\pi)$, where y_0 is the projection of x_0 on \mathbf{S}^2 (see Fig. 3.1). The following lemma allows us to reduce the study to meridian half-planes.

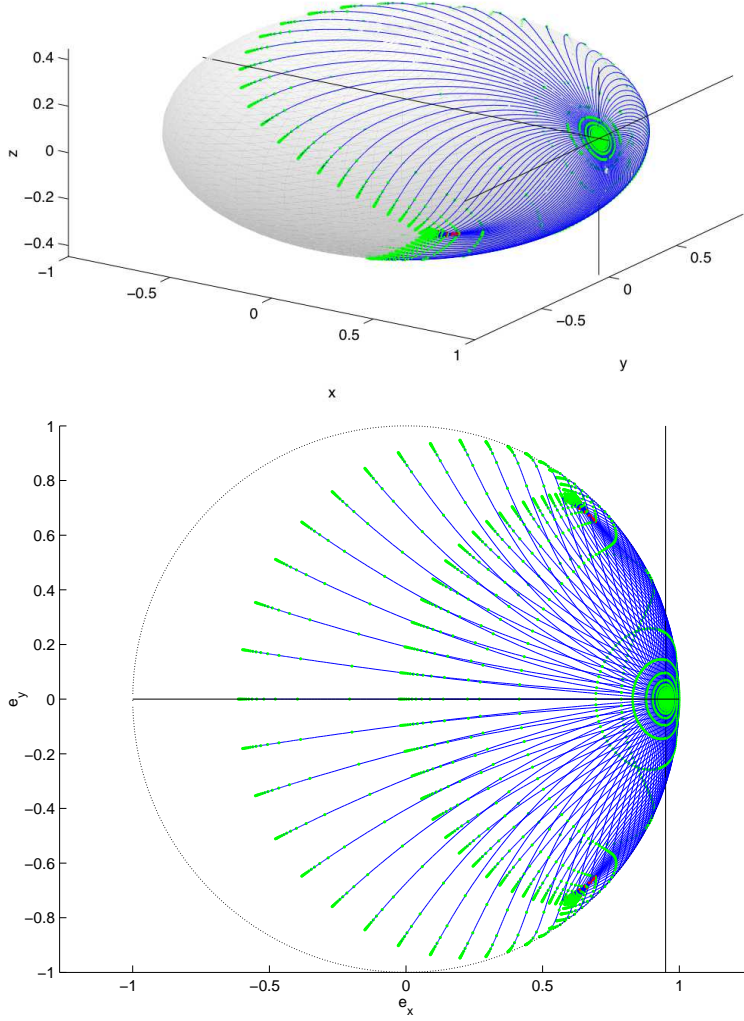


FIG. 3.1. Above, wavefront on the sphere for initial condition $e_0 = 9.5e - 1$ and time $t = \pi$ in Kepler's case ($c = \sqrt{2/5}$). The exponential mapping is surjective if and only if $e_0 < \arcsin(\pi(c - 1/2)) \simeq 4.29e - 1$. Below, projection of this wavefront on the open unit disc, \mathbf{D} . The apparent contour describes the default of surjectivity in the space of ellipses, $X_e = \mathbf{R}_+^* \times \mathbf{D}$, that is before partial compactification.

LEMMA 3.4. *The longest geodesic from any point on \mathbf{S}^2 to the equator is the negative half-meridian, $p_\varphi = -1$.*

Proof. The quadrature (2.3) tells us that, for nonnegative p_{φ_0} , the equator is reached at time $t = t_1 + T/4$. By symmetry, it takes $t = -t_1 + T/4$ to reach the equator on a geodesic such that $-1 < p_{\varphi_0} < 0$ ($-2t_1 > 0$ translation). This time increases towards $\pi/2 + \varphi_0$ which is precisely the time required to reach $\varphi = \pi/2$ on the negative half-meridian defined on $H_2 = 1/2$ by $p_{\varphi_0} = -1$. \square

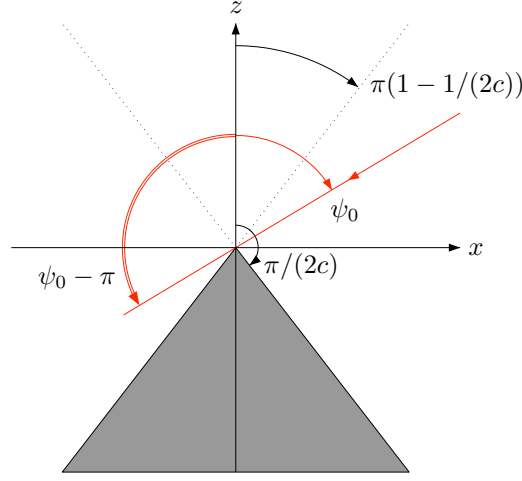


FIG. 3.2. Geodesics in meridian half-planes issuing from a point such that $\psi_0 \geq \pi(1 - 1/(2c))$ —that is $\varphi_0 \geq \pi(c - 1/2)$ —cannot reach points in the sector defined by $\psi \leq \psi_0 - \pi$, and surjectivity is lost.

The picture in meridian half-planes thus characterizes surjectivity, as in Fig. 3.2.

PROPOSITION 3.5. *Given x_0 in X , the exponential mapping \exp_{x_0} is surjective if and only if $\varphi_0 < \pi(c - 1/2)$.*

4. Optimality results. The cut point on a geodesic is the first point where the geodesic ceases to be minimizing, and the first conjugate point is the point where moreover local optimality with respect to neighbouring trajectories is lost. The *conjugate locus* is the set of first conjugate points on geodesics issuing from on given point. Jacobi's theorem, which extends to the more general framework of optimal control [20], asserts that the conjugate locus is a subset of the caustic, subset formed by the first singular values of the exponential mapping. On a complete Riemannian manifold, cut points are either conjugate points, or points on the *separating line* where two geodesics of equal length first intersect [10]. We are thus conducted to examine immersivity and injectivity properties of the exponential mapping. To this end, we begin with a preliminary lifting lemma.

LEMMA 4.1. *Through two points on X which project on \mathbf{S}^2 into points whose distance is less than $c\pi$ passes one geodesic.*

Proof. If the two points project onto the same one on the sphere, they lie on a separatrix in some meridian half-plane, and the result is trivial. Otherwise, let σ denote the distance between them, $0 < \sigma < c\pi$. On the compact manifold (\mathbf{S}^2, g_2) , there is a (minimizing) geodesic of such length joining these points, which is lifted into a (non necessarily minimizing) geodesic on X provided the following system admits at least one solution (p_{r_0}, t) in $] -1, 1[\times \mathbf{R}_+^*$ (see lemma 2.2 and (2.4)),

$$p_{r_0} = ((r_2^2 - r_1^2) - t^2)/(2r_1 t), \quad (4.1)$$

$$t = r_1 \sin(\sigma/c) / \cos(\arcsin(p_{r_0}) + \sigma/c), \quad (4.2)$$

where r_1 and r_2 are the r -coordinates of the two points on X . The (possibly degenerate) hyperbola (4.1) always intersects the second curve in the prescribed domain. Indeed, at $t = r_1$, the point $(r_1, -1)$ belongs to the second curve and is below $p_{r_0}(t = r_1) = (r_2^2 - r_1^2)/(2r_1^2) > -1$. Conversely, when $t \rightarrow \infty$, $p_{r_0} \rightarrow -\infty$ for the first

curve whereas the second one has an horizontal asymptote defined by $p_{r0} = \cos(\sigma/c)$. The two curves must therefore cross somewhere in $] -1, 1[\times \mathbf{R}_+^*$. \square

Given x_0 on X , we denote $\text{Cut}(x_0)$, $C(x_0)$ and $L(x_0)$, the cut locus, conjugate locus and separating line of x_0 , respectively. The same notation is used for y_0 on \mathbf{S}^2 , and we also denote $W(y_0, < t)$ the union of wavefronts from y_0 and times less than t . Let x_0 be a fixed point on X , and y_0 its projection on the sphere.

PROPOSITION 4.2. *A point belongs to the conjugate locus $C(x_0)$ if and only if it projects onto a point in $C(y_0) \cap W(y_0, < c\pi)$.*

Proof. If x belongs to $C(x_0)$, x cannot be on a separatrix in a meridian half-plane starting from x_0 since a direct Jacobi field computation proves that separatrices are without conjugate points. As a result, x projects onto a point y of \mathbf{S}^2 at distance $0 < \sigma < c\pi$ from y_0 . Now, the exponential on X is obtained from the one on the sphere. Namely,

$$\exp_{x_0,t}(p_0) = (r(t, r_0, p_{r0}), \exp_{y_0, \sigma(t, r_0, p_{r0})}((p_{\theta_0}, p_{\varphi_0})/\sqrt{2H_2}))$$

where $\sqrt{2H_2} = (r_0/c)\sqrt{1 - p_{r0}^2}$. A simple rank computation shows that, for positive t and $\sigma(t, r_0, p_{r0})$, $\exp_{x_0,t}$ is singular if and only if $\exp_{y_0, \sigma}$ is singular, too. Thus, y is a conjugate point, and even the first one on the corresponding geodesic (there would be a conjugate point before x , otherwise). Conversely, if y belongs to $C(y_0) \cap W(y_0, < c/\pi)$, there is a geodesic of length $0 < \sigma < c\pi$ joining y_0 and y on which y is the first conjugate point. This geodesic is lifted to a geodesic on X between x_0 and x thanks to Lemma 4.1. The previous rank computation entails that x is a conjugate point of this geodesic, and that there cannot be conjugate points before since this would contradict the fact that y belongs to the conjugate locus. \square

PROPOSITION 4.3. *A point belongs to the separating line $L(x_0)$ if and only if it projects onto a point in $L(y_0) \cap W(y_0, < c\pi)$.*

Proof. If x belongs to $L(x_0)$, there are two geodesics issuing from x_0 , γ_1 and γ_2 , first intersecting at a positive time \bar{t} . Then $r_1(\bar{t}) = r_2(\bar{t})$, which implies $r_1 \equiv r_2$ by virtue of Lemma 2.2 (parabolas of same curvature intersecting at two distinct points are identical). The two geodesics share the same p_{r0} and so project into geodesics of same length on the sphere, $0 < \sigma(\bar{t}, r_0, p_{r0}) < c\pi$. These geodesics first intersect at the projection y of x (since an intersection before would be lifted into an intersection prior to x , otherwise) and y lies in $L(y_0) \cap W(y_0, c\pi)$. Conversely, if y belongs to $L(y_0) \cap W(y_0, c\pi)$, the two geodesics of same length $0 < \sigma < c\pi$ first intersecting there are lifted by Lemma 4.1 into distinct geodesics on X starting from x_0 and first intersecting at x , which is thus on the separating line of x_0 . \square

On (\mathbf{S}^2, g_2) , cut loci are obtained as the closure of separating lines. Since the metric is analytic, the cut locus of any point is a finite tree whose extremities are singularities of the conjugate locus [13, 18]. We now give a complete description of these sets which turn to be completely similar to the cut loci on an oblate ellipsoid of revolution (see Proposition 1.2).

PROPOSITION 4.4. *The cut locus of point of eccentricity $e_0 = \sin \varphi_0$ on the sphere is the line $[\theta_l, 2\pi - \theta_l]$, with $\theta_l = \pi(1 - (1 - \mu^2)e_0)$, included in the antipodal parallel of the point, $\varphi_l = \pi - \varphi_0$. The distance from the point to its cut locus is $t_l = \pi\sqrt{1 - (1 - \mu^2)e_0^2}$, and is reached at the cut point on the corresponding pseudo-equator.*

Proof. In accordance with discrete symmetries discussed in §2, intersecting geodesics of same length are obtained using s_1 or s_2 . Clearly, intersections generated by s_2 come prior to those generated by s_1 (whose length is not less than $T/2$, T being the

period). These intersections are located in the antipodal parallel, and a direct computation on the quadratures of Proposition 2.3 shows that the two geodesics defined on $H_2 = 1/2$ by a fixed p_θ and $\pm p_{\varphi_0}$ intersect at $t_l = T/2 = \pi/\sqrt{1 + (1 - \mu^2)p_\theta^2}$. To conclude, it is sufficient to check that these two geodesics actually do not intersect before. This is obvious since the second one is obtained by a $-2t_1$ translation of the first, with $-2t_1 \leq T/2$, and since φ is monotonic on the half-period $[t_1, t_1 + T/2]$ for nonnegative p_{φ_0} . The closure of the separating line is obtained letting p_θ^2 tend to $G(\varphi_0)$ (letting p_θ tend to 0, one gets for $t_l = \pi$ the intersection of half-meridians oppositely orientated located at $\theta = \pi$ which defines the center of the locus), that is for the associated pseudo-equator, so that $t_l = \pi\sqrt{1 - (1 - \mu^2)e_0^2}$, $\theta_l = \pi(1 - (1 - \mu^2)e_0)$. \square

As a result, we retrieve again the estimation of the injectivity radius of Theorem 2.6. Indeed,

$$i(\mathbf{S}^2) = \inf_{\varphi_0} \pi \sqrt{1 - (1 - \mu^2) \sin^2 \varphi_0} = \mu\pi,$$

reached on the equator where the cut point is a conjugate point. This is in fact the case for any initial point, as a consequence of the following description of the conjugate locus.

PROPOSITION 4.5. *Conjugate times on a geodesic issuing from φ_0 defined on $H_2 = 1/2$ by $0 < p_\theta^2 < G(\varphi_0)$ and positive p_{φ_0} are zeros of*

$$(\cos \varphi_0 / (\sin^2 \varphi_0 - b)^{1/2} - a^3(1 - b)(1 - \mu^2)t/8) \sin(a(t - t_1)) - \cos(a(t - t_1)) = 1 \quad (4.3)$$

with a , b and t_1 functions of p_θ as defined by Proposition 2.3. The distance from any point to its locus is attained by a conjugate point on the corresponding pseudo-equator.

Proof. The constant adjoint state p_θ parameterizes $\Lambda_0 \simeq \S(T_{y_0}^* \mathbf{S}^2)$ whenever p_φ is nonzero. A straightforward differentiation of the quadratures with respect to p_θ in such cases gives the result. Letting p_θ^2 tend to $G(\varphi_0)$, (4.3) degenerates in $\sin(at) = 0$ whose first admissible zero is $T/2$, which is thus the cut and first conjugate point on the associated pseudo-equator. \square

An example of cut and conjugate loci is represented Fig. 4.1 in Kepler's case. The conjugate locus is a deformation of the astroid obtained in the case of the oblate ellipsoid of revolution to which (\mathbf{S}^2, g_2) is conformal.

COROLLARY 4.6. *A necessary and sufficient condition for a geodesic issuing from a point on X with eccentricity e_0 not to have conjugate points or points on the separating line is $e_0 \leq \sqrt{(1 - c^2)/(1 - \mu^2)}$. In Kepler's case, the condition is $\varphi_0 \leq \pi/3$.*

Proof. Simply write that the distance from the projection of the point on \mathbf{S}^2 to its cut locus (see Proposition 4.4) is not less than $c\pi$. \square

We end the paper going back to the original non-compactified averaged metric (0.2) on the space of ellipses, $X_e = \mathbf{R}_+^* \times \mathbf{D}$. The last result essentially asserts that Kepler's geodesics may only lose optimality because of completeness—that is existence—issues.

THEOREM 4.7. *Conjugate loci and separating lines of the averaged Kepler metric on the space of ellipses are always empty.*

Proof. According to the previous analysis, conjugate points or separating points define cut points in projection on the sphere. Now, cut loci on \mathbf{S}^2 are included in antipodal parallels, so that cut points can only be reached by crossing the equator,

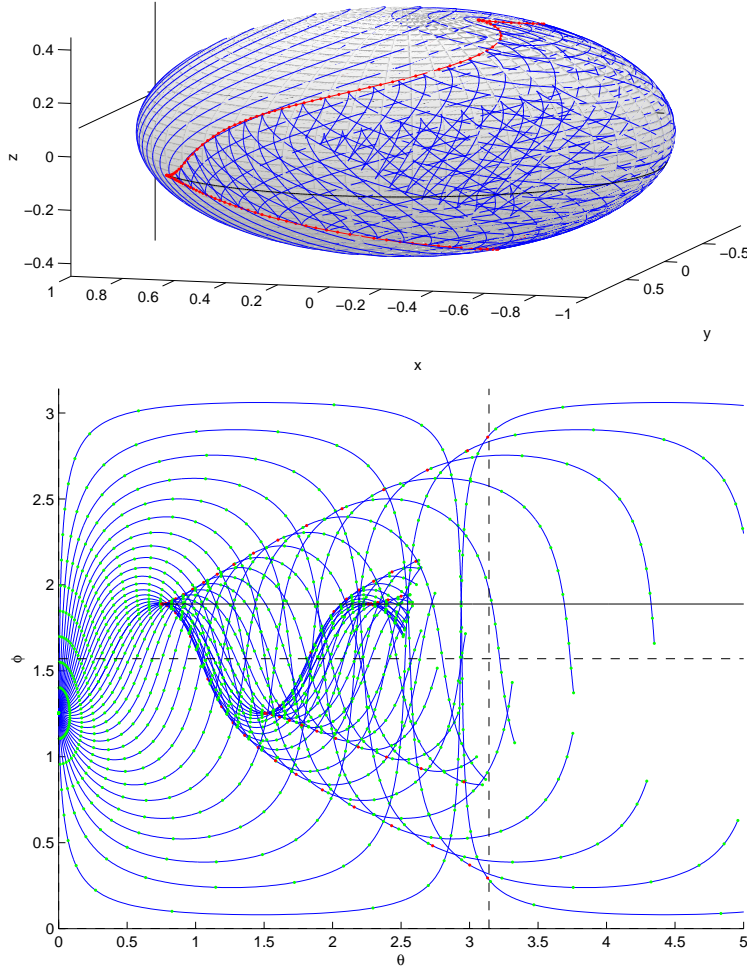


FIG. 4.1. Wavefront, conjugate and cut loci for $e_0 = 9.5e-1$. The bifurcation of the wavefront is observed on (\mathbf{S}^2, g_2) —realized as the oblate ellipsoid to which it is conformal—, and the swallowtail singularities appearing run along the caustic containing the conjugate locus. The inclusion of the cut locus in the antipodal parallel is clearly illustrated in (θ, φ) coordinates.

that is the parabolic boundary $e = 1$ of X_e . There are no such geodesics in the space of ellipses. \square

REFERENCES

- [1] V. I. ARNOLD, A. N. VARCHENKO, AND S.M. GUSEIN-ZADE, *Singularities of differentiable maps*, Vol. 1, Birkhäuser, Boston, 1988.
- [2] J. T. BETTS AND S. O. ERB, *Optimal low thrust trajectories to the Moon*, SIAM J. Applied Dyn. Syst., 2 (2003), pp. 144–170.
- [3] A. BOLSINOV AND A. FOMENKO, *Integrable geodesic flows on two-dimensional surfaces*, Kluwer, New-York, 2000.
- [4] B. BONNARD AND J.-B. CAILLAU, *Riemannian metric of the averaged energy minimization problem in orbital transfer with low thrust*, An. I. H. Poincaré, to appear.

- [5] B. BONNARD, J.-B. CAILLAU, AND E. TRÉLAT, *Geometric optimal control of elliptic Keplerian orbits*, Discrete and Continuous Dynamical Systems, Series B, 5 (2005), pp. 929–956.
- [6] B. BONNARD, J.-B. CAILLAU, AND E. TRÉLAT, *Second order optimality conditions in the smooth case and applications in optimal control*, ESAIM Control, Optimisation and Calculus of Variations, to appear (see also apo.enseeiht.fr/cotcot).
- [7] J.-B. CAILLAU, J. GERGAUD, AND J. NOAILLES, *3D geosynchronous transfer of a satellite: Continuation on the thrust*, JOTA, 118 (2003), pp. 541–565.
- [8] F. R. CHANG-DIAZ, M. M. HSU, E. BRADEN, I. JOHNSON, AND T. F. YANG, *Rapid Mars transits with exhaust-modulated plasma propulsion*, Technical paper 3539, NASA, 1995.
- [9] G. DARBOUX, *Leçons sur la théorie générale des surfaces*, Tome III, Gauthiers-Villars, Paris, 1914.
- [10] M. P. DO CARMO, *Riemannian geometry*, Birkhäuser, Boston, 1992.
- [11] R. EPENOV AND S. GEFFROY, *Optimal low-thrust transfers with constraints: generalization of averaging techniques*, Acta Astronautica, 41 (1997), pp. 133–149.
- [12] J. GERGAUD, T. HABERKORN, AND P. MARTINON, *Low thrust minimum-fuel orbital transfer: a homotopic approach*, J. of Guidance, Control and Dynamics, 27 (2004), pp. 1046–1060.
- [13] H. GLUCK AND D. SINGER, *Scattering of geodesic fields I*, Annals of Math., 108 (1978), pp. 347–372.
- [14] W. P. A. KLINGENBERG, *Poincaré’s closed geodesic on a convex surface*, Trans. AMS, 356 (2004), pp. 2545–2556.
- [15] J. MILNOR, *On the geometry of the Kepler problem*, The American Mathematical Monthly, 90 (1983), pp. 353–365.
- [16] T. NAGELL, *Introduction to number theory*, Wiley, New-York, 1951.
- [17] H.-J. OBERLE AND K. TAUBERT, *Existence and multiple solutions of the minimum-fuel orbit transfer problem*, JOTA, 95 (1997), pp. 243–262.
- [18] H. POINCARÉ, *Sur les lignes géodésiques des surfaces convexes*, Trans. AMS, 5 (1905), pp. 237–274.
- [19] Y. L. SACHKOV, *Discrete symmetries in the generalized Dido problem*, Sbornik Math, 197 (2006), pp. 235–257.
- [20] A. V. SARYCHEV, *The index of second variation of a control system*, Math USSR Sbornik, 41 (1982), pp. 338–401.