Qualitative Theory of Dynamical Systems (2020) 19:31
https://doi.org/10.1007/s12346-020-00363-8

Minimizers for the Kepler Problem

Richard Montgomery[1]

Received: 6 August 2019 / Accepted: 9 October 2019 / Published online: 31 January 2020

Abstract
We characterize the minimizing geodesics for the Kepler problem endowed with the Jacobi-Maupertuis metric. We focus on the positive energy case, but do all energies. The more complicated negative energy case was solved in Jacobi (Crelles J 17:68–82, 1837. https://doi.org/10.1515/crll.1837.17.68), with his work translated and completed by Todhunter (Researches in the Calculus of Variations, Principally on the Theory of Discontinuous Solutions. Macmillan and Co., Cambridge, 1871), and later summarized in Wintner's book. Our discussion of these old results includes a new proof for the positive energy case and perspectives coming from metric and differential geometry. For the negative energy result we need Lambert's theorem which we discuss.

Keywords Jacobi–Maupertuis metric · Minimizing geodesics · Lambert's theorem · Kepler's problem

1 Introduction and Overview

Solutions of the planar Kepler problem

$$\ddot{q} = -q/\|q\|^3, q \in \mathbb{R}^2$$

lie on conics one of whose foci is the origin. Following Albouy [1] , by a "Keplerian arc" we will mean a finite arc $q([a, b])$ of such a solution. These include the rectilinear arcs which lie on degenerate conics: rays or line segments having one endpoint the origin.

The energy

$$H = \frac{1}{2}|\dot{q}|^2 - \frac{1}{|q|}$$

✉ Richard Montgomery
rmont@ucsc.edu

1 Mathematics Department, University of California, Santa Cruz, CA 95064, USA

and angular momentum

$$J = q \wedge \dot{q}$$

are constant along Keplerian arcs. The arc is rectilinear if and only if $J = 0$. A Keplerian arc with energy $H = h$ is, upon reparameterization, a geodesic for the *Jacobi-Maupertuis* [JM] metric

$$ds_h^2 = 2\left(h + \frac{1}{|q|}\right)|dq|^2.$$

defined on the Hill region $\{q : h + 1/|q| \geq 0\} \subset \mathbb{R}^2$. As such, a Kepler arc of energy h must extremize the JM length $\ell_h(c) = \int_c ds_h$ among all curves sharing its endpoints, but may fail to minimize this length.

Question 1 *Among the Keplerian arcs of energy h, which are minimizing geodesics?*

Theorem 1.1 *For positive energy, a non-rectilinear Keplerian arc with endpoints A, B minimizes the JM length among all curves sharing its endpoints if and only if the origin is not in the interior of the convex region bounded by the arc and the chord AB.*

Definition 1.1 We call the arc "direct" if the origin is not in the interior of its convex hull and "indirect" otherwise.

With this terminology, the theorem asserts that direct hyperbolic arcs minimize while indirect ones do not.

1.1 History. Disclaimer. Acknowledgements. Memorium

Theorem 1.1 can be found in section 256 of Wintner's classic book [11], while sections 254 and 257 deal with the negative and zero energy versions of this theorem. The negative energy version, Theorem 4.1 below, appears to be due to Jacobi [3], and was translated by Todhunter [10], section 226, (p 251) which is quite likely where Wintner learned of it. Jacobi's description of the minimizers did not include what we call "turnpike paths", described in Fig. 4 below, which are paths which make use of the Hill boundary. Todhunter investigated these alternate, often shorter paths, under the name 'discontinuous solutions" in his book [9]. This book contains further details around Jacobi's work in sections 179 to 183.

With all this previous work, why restate and reprove the theorems? First, the simple reflection based proof of half of Theorem 1.1 found below in Sect. 2.1 is not found in these earlier works. Second, the complete understanding of the minimizers for the Kepler problem deserves to be widely known. Third, our metric space analysis perspective on the Jacobi-Maupertuis metric is not in these earlier works. Fourth, within this general metric context, the question "what are the infinite minimizers?" is of current interest, and impacts the weak KAM theory as applied to the general dynamics of the non-negative energy n-body problem. See for example [4] and [5] and compare results there with Corollary 3.1 below.

In memory. I would like to express my condolences to the students and family of Florin Diacu, as well as my thanks to him. Through the force of his personality, his writing, and his organizational efforts he supported and encouraged many of us working in mathematical celestial mechanics. He had a particular soft spot for the N-body problem on curved backgrounds, and I hope that it would have given him an amused smiled to reflect on the Kepler problem as defining a curved substructure for the Euclidean plane.

2 Proving the Theorem

2.1 Cut Points: Proving Half of Theorem 1.1

We prove that if the arc is indirect then it fails to minimize. So assume that the origin is in the interior of the convex hull of the arc. Parameterize the arc by a solution $B(t)$ with $B(0) = A$ and $B(t_1) = B$ for some $t_1 > 0$. Consider the moving chords $AB(t)$. Since the origin is in the interior of the convex hull , there will be a t_*, $0 < t_* < t_1$ such that this chord $AB(t_*)$passes through the origin. Write $B_* = B(t_*)$ (Figs. 1, 2).

We claim that B_* is a cut point along $B(t)$. First, recall the notion of 'cut point' along a geodesic σ in Riemannian geometry. As we leave the initial point $A = \sigma(0)$ of the geodesic we ask : does the arc $\sigma([0, t])$ minimize the JM length between A and $\sigma(t)$? It will minimize for all sufficiently short times $t \leq \epsilon$. A "cut point" along the geodesic is a point $B_* = \sigma(t_*)$ such that another distinct geodesic with the same length as $\sigma([0, t_*])$ joins A to B_*. A basic theorem from Riemannian geometery asserts that if B_* is a cut point, then the geodesics $s([0, t])$ fail to minimize for all $t > t_*$.

We apply these considerations to our geodesic, the curve $B(t)$, reparameterized by JM arclength. To see that B_* is a cut point, form the line $\ell = OA$ and note that it passes through B_*. Any linear isometry of the plane is an isometry of the JM metric, and in particular reflection about ℓ is a JM isometry. Applying this reflection to the arc $B([0, t_*])$ joining A to B_* we obtain a distinct geodesic arc (since the original arc is not rectilinear) joining A to B_* and having the same JM length as the original. We have shown that B_* is a cut point along $B(t)$. Since $t_1 > t_*$ our arc from A to $B = B(t_1)$ fails to minimize. □

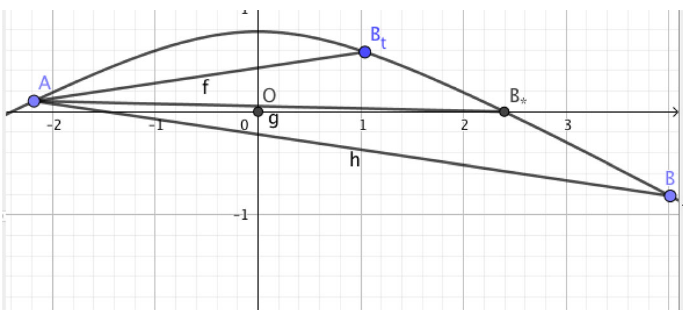

Fig. 1 The moving chord passes through the origin

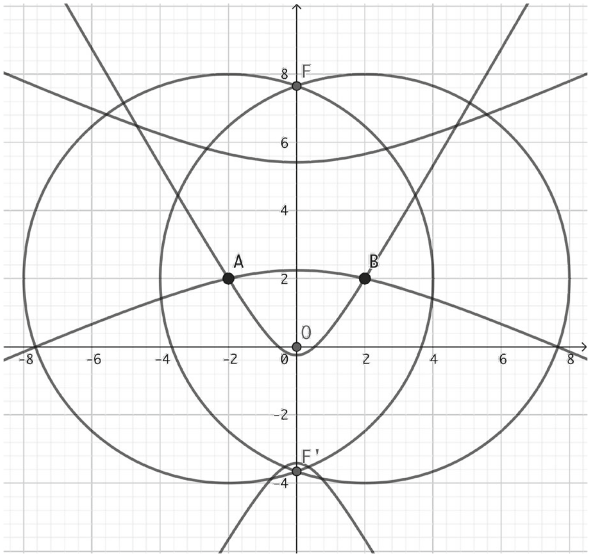

Fig. 2 The Gauss construction of the two hyperbolic branches passing through A and B. The empty foci F and F′ are constructed by intersecting circles of radius $r_A = OA + 2a$ and $r_B = OB + 2a$ about A and B. The branch for F is direct and for F' is indirect

2.2 A Method of Gauss: Just Two Hyperbolas

To prove the other half of the Theorem 1.1 we use a lemma which is the positive energy version of a method of Gauss explained in remark 3, section 2 of [1]. See also Wintner [11].

Lemma 2.1 *Let $h = 1/2a > 0$ and suppose that $A, B \in \mathbb{R}^2$ are not collinear with 0. Then exactly two Kepler arcs with energy h pass through A and B. On of these arcs is direct, and the other is indirect.*

If $A, 0, B$ are collinear with 0 between A and B then exactly two Kepler arcs of energy h connect A to B. These arcs are related by reflection about line $A0 = OB$.

If $A, B, 0$ are collinear with B between A and 0 or with A between B and 0 then exactly one branch of energy h passes through A and B. This arc is rectilinear.

Proof The proof we present is similar to that found in the 2nd paragraph of section 256 of [11]. If A, B are not collinear with 0 then any branch passing through them is a hyperbola and not a ray. Hyperbolae have two foci. One of ours is 0. The location of the other "empty focus" F, together with the value of the energy $h = 1/2a$, determines the Kepler branch of such a hyperbola as the locus of points X such that

$$|XF| - |X0| = 2a, h = 1/2a$$

(The distance between F and 0 is $2ea$ where $e > 1$ is the eccentricity. See exercise 4.1 and formulae in sections 1.4 and 1.5 of [8]) Since both A and B lie on this alleged hyperbola we have

$$|AF| - |A0| = 2a, \qquad |BF| - |B0| = 2a \tag{1}$$

We view these as equations for F. Rewritten as $|AF| = 2a + |OA|$ and $|BF| = 2a + |OB|$ they say that F must lie simultaneously on the circle of radius $r_A := 2a + |OA|$ about A and the circle of radius $r_B := 2a + |OB|$ about B. Now two circles in the Euclidean plane intersect in two points if and only if the sum of their radii is greater than the distance between their centers and neither circle is contained within the other. In terms of our variables these inequalities, necessary for there to be two intersections, are

$$r_A + r_B > |AB|, r_B < |AB| + r_A, \text{ and } r_A < |AB| + r_B. \tag{2}$$

In our case $r_A + r_B = 4a + |OA| + |OB| \geq 4a + |AB| > |AB|$ while $|AB| + r_A = |AB| + |OA| + 2a \geq |OB| + 2a = r_B$ and similarly $|AB| + r_B \geq r_A$. This shows that the first inequality of the three of inequalties (2)always holds in its stated strict form, while the remaining two inequalities may not hold in their strict form but instead may be equalities. If A, B, O are not collinear then all inequalities in the above reasoning are strict so that all inequalities of (2) hold and the two circles intersect in exactly two points, namely our two foci, call them F and F'. One of these, say F, lies on the opposite side of the line AB as 0, while the other focus F' lies on the same side of AB as O. One checks that the arc associated to F is direct while the arc associated to F' in indirect.

If A, O, B are collinear with O between A and B then the only possible Keplerian branches connecting them are again hyperbolas. We again need only check all three inequalities of (2) are strict to insure exactly two such arcs. We saw that the first inequality always holds in the strict sense. For the second and third use that $|AB| > |OA|, |OB|$ so that $|AB| + r_A = |AB| + |OA| + 2a > |AB| + 2a > |OB| + 2a = r_B$. Similarly $|AB| + r_B > r_A$. Thus the two circles intersect in two points, our two foci again. As per the argument of the first half, the hyperbolas defined by these two foci are related by reflection.

The rectilinear case arises if and only if $OA + AB = OB$ or $OB + AB = OA$ which is equivalent to $r_B = |AB| + r_A$ or to $r_A = |AB| + r_B$, the case where one circle contains the other, with the two intersecting tangentially at O. □

2.3 The Other Half of the Proof of Theorem 1.1

We prove that if the arc is direct then it minimizes the JM length among all curves joining A to B.

There is a JM minimizer joining A to B, by standard results in the calculus of variations. Moreover this minimizer cannot contain O in the interior of its arc by the Maupertuis version of Marchall's lemma. See Lemma 1 of [6]. It follows that the only curves competing to be minimizers are the Keplerian arcs. We must show our arc is the shortest among all such arcs. But by Lemma 2.1 there is at most one other competing arc!

If the endpoints A, B are not collinear with O then, by that lemma, there is exactly one other arc through A and B having energy h. This other arc is indirect and consequently, by the first half of the theorem, cannot minimize. So our branch is the minimizer.

If A, B are collinear with O with O between them, then, according to the lemma the two branches are related by reflection about the line AB and their energies and JM lengths are the same. Both are minimizers. This is the cut point case in the proof of Sect. 2.1.

If A, B are collinear with either B between A and O or A between O and B then there is a unique branch and it is rectilinear. Necessarily it is the minimizer. (This minimality can also be established directly using $dr^2 \leq dr^2 + r^2 d\theta^2$.) □

3 Motivation and Metric Spaces

The seed for this paper was the desire to characterize the global minimizers for the positive energy case. When $h \geq 0$ the JM Riemannian metric ds_h^2, defined on the entire plane minus the origin, extends to the origin as a metric, where 'metric' now means distance function. Define the "distance" between two points of the plane $\mathbb{R}^2$ to be the infimum of the JM lengths $\int ds_h$ among all curves joining them. The rectilinear arc $[A, 0]$ from a point $A \in \mathbb{R}^2$ in the plane to the origin 0 of the plane has finite JM length and hence 0 is a finite distance from A. In this way we get a distance function on the entire plane, one which gives it the structure of a complete metric space, Riemannian everywhere except at 0. At the collision point 0 the metric enjoys a kind of conical singularity. On any metric space we have the notion of "length of a path", "minimizing geodesics" and "geodesic". See [2]. On $\mathbb{R}^2$, endowed with the JM distance function just described, the geodesics which do not pass through the origin are the Keplerian arcs for this energy. Now let M be a metric space, $J \subset \mathbb{R}$ an infinite interval, and $\sigma : J \to M$ a curve parameterized by arclength. Then σ is called a 'global minimizer", or synonymously, an "infinite minimizing geodesic", if its restriction to any compact subinterval $[a, b] \subset J$ is a minimizing geodesic between its endpoints $\sigma(a)$ and $\sigma(b)$. We say that a global minimizer σ is a "maximal global minimizer" if it is impossible to extend it to a larger interval $J' \supset J$ in such a way that the extended curve $s' : J' \to M$ remains a global minimizer.

Corollary 3.1 *Let $\sigma : \mathbb{R} \to \mathbb{R}^2$ be a Kepler hyperbola of energy h. Write L for the line through the origin parallel to the positive time asymptote of the hyperbola, and let $A_* = \sigma(a_*)$ denote the point of intersection between L and the hyperbola. Then $\sigma([a_*, \infty))$ is a maximal global JM minimizer. (See* Fig. 3 *for an illustration of this construction.)*

Similarly, let L_- be the line through the origin parallel to the the negative time asymptote, and $\sigma(b_*)$ for the intersection point of L_- with the hyperbola. Then $\sigma((-\infty, b_*])$ is a maximal global JM minimizer.

Remark If t_* is the time of apocenter along the hyperbola in the theorem, which is to say, the time of the hyperbola's closest approach to the origin, then $b_* < t < a_*$.

Proof of Cor. 3.1. We have seen that for finite intervals $[a, b]$, the arc $\sigma([a, b])$ minimizes if and only if the convex hull $C_{a,b}$ of this arc does not contain the origin in its interior. If $[c, d] \supset [a, b]$ then $C_{c,d} \supset C_{a,b}$. Taking limits as $b \to \infty$ one sees

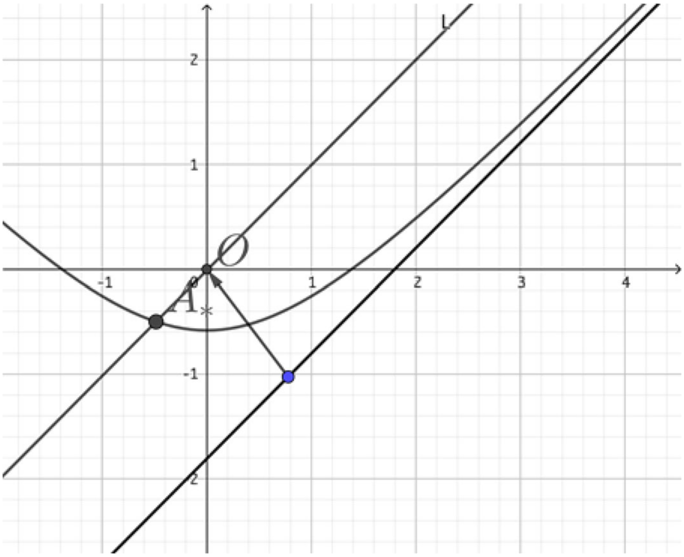

Fig. 3 A Kepler hyperbola. Point $A_* = \sigma(a_*)$ is the endpoint of the maximal global minimizer subarc of this hyperbola, unbounded in the positive direction. A_* is the intersection of the hyperbola with the line L parallel to the hyperbola's positive asymptote and passing through the origin

that the convex hull of $\sigma([a, \infty)$ has boundary consisting of two infinite curves, the arc $\sigma([a, \infty))$ of the hyperbola and the ray leaving $\sigma(a)$ and parallel to L. By the above discussion, we are in the borderline case between minimizing and failing to minimize if and only if this ray passes through the origin. This borderline case is the case corresponding to $a = a_*$ as described in the corollary, the case of the maximal global minimizer.

The case of the past asymptotic maximal global minimizer proceeds similarly. □

Remark 3.1 The corollary continues to hold in the parabolic case, $h = 0$ with essentially the same proof. A quite different proof, based on an isometry from the JM metric $ds^2_{h=0}$ to that of a cone over a circle of radius $1/2$, can be found in [7].

4 The Negative Energy Case

If $h = -1/2a < 0$ then the JM metric is defined on the Hill region $D_h := \{q : h + 1/|q| \geq 0\}$ which is the disc of radius $2a$ centered at the origin. The infinitesimal distance ds_h vanishes on the boundary of this region, which we call the Hill boundary, and which equals the Euclidean circle of radius $2a$ centered at the origin. As a result of this vanishing, any curve lying on the Hill boundary has zero JM length, and hence the JM distance between any two points on the Hill boundary is zero. Consequently, the induced metric distance function on the Hill region is not a true distance function, but rather a pseudo-distance. There is a new type of minimizer which takes advantage of this zero-cost motion along the Hill boundary. See Fig. 4. Such a minimizer connects

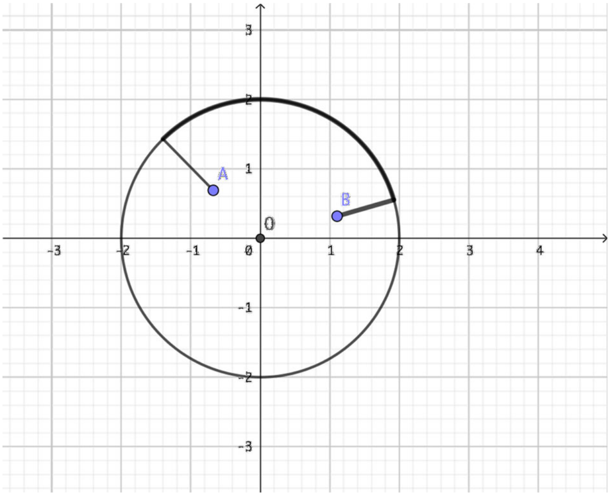

Fig. 4 A turnpike path from A and B. Here $a = 1$

points A and B by travelling radially out from A to the Hill boundary, angularly along the Hill boundary until it reaches the ray through B, and then back inwards radially to B. We will borrow terminology from economics and refer to these new minimizers as "turnpike paths". Todhunter called them "discontinuous minimizers" and investigated them extensively towards the end of chapter 8 of [9].

Theorem 4.1 *For negative energy, a non-rectilinear Keplerian arc with endpoints A, B and lying on an ellipse $\mathcal{E}$ minimizes JM length* **among Keplerian arcs** *if and only if the convex region bounded by the arc and the chord AB does not contain either the origin O nor the empty focus F of the ellipse $\mathcal{E}$ in its interior.*

However, it may happen that the minimizer from A to B is a turnpike path. This possibility does happen if A and B form the semi-major axis of the ellipse, or any chord sufficiently close to this axis, i.e, provided that $|AB| = 2a - \epsilon$ for $\epsilon = 0$ or ϵ sufficiently small, where $h = -1/2a$ is the energy.

A Proof of the Half of Theorem 4.1. The proof for $h > 0$ in 2.1 holds verbatim for $h \leq 0$ and shows that an arc whose convex hull contains 0 in its interior fails to minimize. Paragraph 3 of section 253 of [11] shows that if $AB(t_*)$ passes through F then t_* is a conjugate point for $B(t)$, and hence if F is in the interior of the convex hull the arc fails to minimize.

To show that the arc minimizes among Keplerian arcs if O and F are exterior to the convex hull described, use the fact that there is at exactly one other Keplerian arc joining A and B and having energy h. This fact, the elliptic version of Lemma 2.1, is found in [9] or [11] as above. In addition, if F' denotes the empty focus of this other Keplerian arc, then the convex hull of this other arc will contain O or F' in its interior and hence this other arc does not minimize.

In case F or O lies on the boundary of the convex hull of the given arc we proceed by cases as follows. If F on the boundary then the Keplerian arc from A to B with energy h is unique, and B is the first conjugate point along this arc as it leaves A. If O on the boundary then, by the reflection principle, as in Sect. 2.1, the other arc is the reflection of the given arc along line AB, it has the same length as the given arc, and B is a cut point along the given arc as it leaves A. □

We will prove the second half of Theorem 4.1 in the next section.

5 Lambert to the Rescue!

We describe Lambert's theorem, following Albouy [1]. The theorem turns the rectilinear paths into measuring sticks for JM lengths yielding a proof of the second half of Theorem 4.1, as well as alternative proofs of all of Theorems 1.1 and 4.1.

We begin with an artifice for extending rectilinear arcs through collision. As a rectilinear solution $q(t)$ approaches collision, let us agree to continue it through collision by reflecting it off of collision, insisting that it remain a solution, having the same energy as it had before collision and remaining on the same ray. Thus, if $q(t)$ is defined for $a < t < 0$ and $lim_{t\to 0} q(t) = 0$, we set $q(0) = 0$ and $q(t) = q(-t)$ for $t > 0$ small. In this way any rectilinear solution uniquely extends so as to define a map $q : \mathbb{R} \to \mathbb{R}^2$ defined for all time and whose image lies on a single ray. We continue to call such curves "Keplerian conics" and the restriction of such a curve to any finite time interval $[a, b]$ a "Keplerian arc".

Remark 5.1 A rectilinear conic having energy $h \geq 0$ has exactly one collision and tends to infinity in both backwards and forwards time. A rectilinear conic of energy $h = -1/2a < 0$ sweeps out a radial interval of length $2a$ with one endpoint collision, the other on the 'Hill boundary' $|q| = 2a$, oscillating periodically between these extremes, taking half a Keplerian period, which is to say, a time $\pi a^{3/2}$ to perform the transit from one endpoint to the other.

Remark 5.2 The C^0–closure of the space of non-rectilinear Keplerian arcs consists of the rectilinear Keplerian arcs in our new extended sense, together with the space of non-rectilinear Keplerian arcs.

Time translation by $t_0 \in \mathbb{R}$ acts on the space of collinear arcs by sending a solution arc or extended solution arc $q(t)$ defined on the interval $[a, b]$ to the arc $t \mapsto q(t - t_0)$ defined on the interval $[a + t_0, b + t_0]$.

Definition 5.1 We write $\mathcal{E}$ for the space of Keplerian arcs, modulo time translation, with rectilinear arcs, extended as described above through collision, included.

$\mathcal{E}$ is a 5-dimensional smooth manifold. One way to parameterize it is to insist that the time intervals $[a, b]$ parameterizing the arcs starts at $a = 0$. Then we can use the initial conditions A, v_A at the initial time $t = 0$ together with the time of flight $\Delta t = b - a = b$ as coordinates. The extended retilinear arcs correspond to inital conditions where the velocity v_A is parallel to the position A and the flight time Δt is greater than the time to collision. The two endpoints $A = q(0)$, $B = q(b)$, the energy H, the time-of-flight $b = \Delta t$ and the JM length $w = \int_a^b ds_{H=h}$ are all smooth functions on $\mathcal{E}$.

Theorem 5.1 *The JM length w and time of flight Δt is constant along any continuous path in the space $\mathcal{E}$ of Keplerian arcs along which the values of $|OA| + |OB|$, $|AB|$ and H are constant.*

Every Keplerian arc $\mathcal{K} \in \mathcal{E}$ can be connected to a rectilinear arc $\mathcal{R} \in \mathcal{E}$ by a continuous path $\lambda \mapsto \mathcal{K}(\lambda)$ of Keplerian arcs along which $|OA| + |OB|$, $|AB|$ and H are constant.

Proof See [1] or references therein.

When $\mathcal{R}$ and $\mathcal{K}$ are as in the second half of the theorem, we will call $\mathcal{R}$ a "rectilinear representative of $\mathcal{K}$". The group of rotations acts on $\mathcal{E}$, preserving $|OA|+|OB|$, $|AB|$, and H. Using a rotation we can bring any rectilinear representative so as to lie on the non-negative x-axis, $y = 0, x \geq 0$. We will refer to such a rectilinear arc $\mathcal{R}$ as a 'standard arc", or a "standard representative" of $\mathcal{K}$. A standard arc at energy h must lie in the intersection of the x-axis with the Hill region. If $h \geq 0$ this provides no constraint: the arc can lie anywhere within the ray. If $h = -1/2a < 0$ then the arc must lie within the interval $0 \leq x \leq 2a$. The length of a standard arc is measured using the restriction of the JM metric to the axis, which is to say, by integrating $(\sqrt{1/x + h})dx$ along the arc.

Definition 5.2 By the *standard model* for energy h we will mean the corresponding domain, (interval $[0, 2a]$ for $h = -1/2a$, or the or entire axis $[0, \infty)$ for $h \geq 0$) af the x-axis, endowed with the metric $(\sqrt{1/x + h})dx$.

If the endpoints A, B of the Keplerian arc $\mathcal{K}$ are distinct then there are exactly two possible endpoints for any of its standard representatives $\mathcal{R}$. To see this, set

$$S = |OA| + |OB|$$
$$C = |AB|$$

and let $x_1, x_2 \geq 0$ be possible endpoints. Then $x_1 + x_2 = S$ while $|x_1 - x_2| = C$. Now the map $(x_1, x_2) \mapsto (x_1 + x_2, x_1 - x_2)$ is an invertible linear map. If $A \neq B$ we have $C \neq 0$. It follows that there are exactly two solutions to our system of equations, namely (x_1, x_2) and (x_2, x_1) with $x_1 = (1/2)(S + C)$, $x_2 = (1/2)(S - C)$.

5.1 Proving the Rest of Theorem 4.1

Take AB to be the semi-major axis of a Keplerian ellipse of energy $h = -1/2a$. We will show that the length of the corresponding half-ellipse is strictly greater than the JM length of the turnpike solution connecting A to B.

First we compute the JM length of the half-ellipse to be $\pi\sqrt{a}$. To do this, observe that for the semi-major axis we have that $C = S = 2a$ which are the same values that C and S have for the half-circle of radius a, which is a Keplerian arc of energy h and endpoints a diameter of this circle. On the semicircle, in polar coordinates, we have that $ds_h = \sqrt{a}d\theta$. Integrating from 0 to π yields the claim.

The length of the turnpike path is $\int_{|0A|}^{2a} f(x)dx + \int_{|0B|}^{2a} f(x)dx$ where $f(x) = \sqrt{2(1/x + h)}$. Since $f(x)$ is monotone decreasing in x , we have $\int_{x_0}^{x_0+\delta} f(x)dx <$

$\int_0^\delta f(x)ds$ for any $x_0, \delta > 0$ such that $[x_0, x_0 + \delta] \subset [0, 2a]$. Now use $2a = |OA| + |OB|$ and take $x_0 = |0B|, \delta = |OA|$ to get that $\int_{|0B|}^{2a} f(x)dx < \int_0^{|OA|} f(x)dx$. Thus $\int_{|0B|}^{2a} f(x)dx + \int_{|0A|}^{2a} f(x)dx < \int_0^{|OA|} f(x)dx + \int_{|0A|}^{2a} f(x)dx = \int_0^{2a} f(x)dx$. Now this last integral equals $\pi\sqrt{a}$ since its endpoints, $0, 2a$ are those of the standard representative of the half-circle, $C = S = 2a$. Thus we get $\int_{|0A|}^{2a} f(x)dx + \int_{|0B|}^{2a} f(x)dx < \pi\sqrt{a}$ □

5.2 Alternate Proofs of Theorem 1.1 and 4.1

We finish by showing how to use Lambert's theorem to give alternative proofs of these theorems.

Parameterize the energy h conic $\mathcal{K}$ by the solution curve $B : \mathbb{R} \to \mathcal{K} \subset \mathbb{R}^2$ starting at A, so that $B(0) = A$. For each $t \geq 0$ let $\mathcal{K}_t \subset \mathcal{K}$ denote the arc from A to $B(t)$ obtained by restricting B to $[0, t]$. We will prove that $\mathcal{K}_t$ minimizes, **among all competing Keplerian arcs**, up until the first time that t_* that the chord $AB(t)$ passes through either O or F. We then make remarks on the turnpike paths and the difference between positive and negative energies.

By a basic theorem in Riemannian geometry the $\mathcal{K}_t$ are global minimizers for t small. Use Theorem 5.1 to form the corresponding standard arc $\mathcal{R}(t)$ with endpoints $x_A(t), X_B(t)$ on the x-axis. Both families $\mathcal{K}_t, \mathcal{R}_t$ depends continuously on t, as do their JM lengths which increase monotonically with t and are equal. The arc $\mathcal{R}(0)$ is a point curve with 0 length and equal endpoints $x_A(0) = x_B(0) = |OA|$. As t increases its endpoints gradually move apart, with the JM length between them continuing to equal the length of $\mathcal{K}_t$. The arcs $\mathcal{R}(t)$ minimize within the standard model as long as they do not retrace themselves, which is to say, as long as they have not bounced off of the collision point or Hill boundary. Thus retracing begins at the instant t_* at which one of the endpoints hits collision $x = 0$, or, only possible in the negative energy case, one endpoint hits the Hill boundary $x = 2a$. The crucial observation is

- (a) one endpoint of the standard arc $\mathcal{R}_t$ is the collision point 0 if and only if chord $AB(t)$ passes through 0.
- (b) one endpoint of the standard arc $\mathcal{R}_t$ is the Hill boundary point of the standard ray if and only if the chord $AB(t)$ passes through the empty focus F.

Proof of (a). Having chord AB pass through the origin is characterized by the equality $|A0| + |B0| = |AB|$. Otherwise $|AO| + |OB| > |AB|$. On the other hand, for a pair of points x_1, x_2 on the standard ray, having one of them equal to the collision point x = 0 is characterized uniquely by the equality $x_1 + x_2 = |x_2 - x_1|$. In terms of chord C and sum S, this equality is $C = S$. Since $\mathcal{R}_t$ and $\mathcal{K}_t$ share the same values of C and S these two events happen simultaneously.

Proof of (b). If the chord passes through the empty focus F we have that $|AB| = |AF|+|FB|$. Then $|OA|+|OB|+|AB| = |OA|+|AF|+|OB|+|BF| = 2a+2a = 4a = -2/h$. On the other hand, if the chord *does not* pass through the empty focus then $|AB| < |AF|+|FB|$ and so $|OA|+|OB|+|AB| < |OA|+|AF|+|OB|+|BF| = -2/h$. This shows that the event "passing through the empty focus" is characterized by the equality $S + C = -2/h$. Now draw the points x_1, x_2 on the standard line segment $[0, 2a]$. A simple bit of exploration shows that $x_1 + x_2 + |x_1 - x_2| \leq 4a$ with equality if and only if one or the other of x_1, x_2 equals the Hill boundary point $x = 2a$. Thus 'hitting the Hill boundary" is characterized in the rectilinear model by the same equality $S + C = -2/h$.

Alternate Proof of Theorem 1.1. Hitting the Hill boundary is impossible. And the only competing paths are other Keplerian arcs. Any other Keplerian arc $\tilde{\mathcal{K}}$ between A and $B = B(t)$ has a rectilinear representative, $\tilde{\mathcal{R}}$ whose endpoints x_A, x_B are the same as those of $\mathcal{R}_t$. There are only two such rectilinear paths connecting two points on the standard ray, one which bounces off collision once, and one which does not. The rectilinear path which does not bounce is the minimizer, and corresponds to the convex hull described in the theorem excluding the origin from its interior. □

Alternate Proof of the First Half of Theorem 4.1. Any Keplerian arc $\tilde{\mathcal{K}}$ between A and $B = B(t)$ has a standard representative, $\tilde{\mathcal{R}}$ whose endpoints x_A, x_B are the same as those of $\mathcal{R}_t$. There are now many such rectilinear paths connecting two points x_A, x_B on the standard ray, depending on how many times a path bounce off collision and the Hill boundary. Only one of these minimizes, the one which does not bounce at all. This standard representative is characterized by the fact that its corresponding Keplerian arc has a convex hull as described in the theorem, i.e. one which excluded the origin and the empty focus from its interior. □

Acknowledgements Thank you to Alain Albouy for his clear writings on Lambert's theorem, for many useful email discussions, and pointers to the literature. Thank you to Hector Sanchez and Rick Moeckel for ongoing email discussions and interest. Thank you to Andrea Venturelli and Ezequiel Maderna for catching a basic error in the statement of Corollary 3.1.

References

1. Albouy, A.: Lambert's Theorem: Geometry or Dynamics? arXiv:1711.03049
2. Burago, D., Burago, Y., Ivanov, S.: A course in metric geometry. Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence, RI, (2001)
3. Jacobi, C.G.: Zur Theorie der Variations-Rechnung und der Differential-Gleichungen, ournal für die reine und angewandte Mathematik. Crelles J. **17**, 68–82 (1837). https://doi.org/10.1515/crll.1837.17.68
4. Maderna, E., Venturelli, A.: Viscosity Solutions and Hyperbolic Motions: A New PDE method for the N-body problem arXiv:1908.09252
5. Moeckel, R., Montgomery, R., Sanchez, H.: Free time minimizers for the planar three-body problem, Celestial Mechanics and Dynamical Astronomy, Vol. 130, no. 3, (2018)
6. Moeckel, R., Montgomery, R., Venturelli, A.: From brake to syzygy. Arch. Rat. Mech. Anal. **204**(3), 1009–1060 (2012)
7. Montgomery, R.: The Kepler Cone and Weak KAM. preprint, available by email request
8. Pollard, H.: Celestial Mechanics , MAA, (1977)
9. Todhunter, M.A.: Researches in the Calculus of Variations, Principally on the Theory of Discontinuous Solutions. Macmillan and Co., Cambridge (1871)
10. Todhunter, M.A.: History of the Calculus of Variations. Macmillan and Co., Cambridge (1871)
11. Wintner, A.: Analytical Foundations of Celestial Mechanics. Princeton Mathematical Series, vol. 5. Princeton U. Press, Princeton (1947)