MINIMAL GEODESICS OF THE ISOSCELES THREE BODY PROBLEM

RICHARD MOECKEL

Abstract. The isosceles three-body problem with nonnegative energy is studied from a variational point of view based on the Jacobi-Maupertuis metric. The solutions are represented by geodesics in the two-dimensional configuration space. Since the metric is singular at collisions, an approach based on the theory of length spaces is used. This provides an alternative to the more familiar approach based on the principle of least action. The emphasis is on the existence and properties of minimal geodesics, that is, shortest curves connecting two points in configuration space. For any two points, even singular points, a minimal geodesic exists and is nonsingular away from the endpoints. For the zero energy case, it is possible to use knowledge of the behavior of the flow on the collision manifold to see that certain solutions must be minimal geodesics. In particular, the geodesic corresponding to the collinear homothetic solution turns out to be a minimal for certain mass ratios.

1. Blown-up Jacobi Coordinates

Consider the planar three-body problem with masses $m_1 = m_2 = 1$, $m_3 > 0$. Then there is an invariant subsystem with two degrees of freedom such that the triangle formed by the bodies remains isosceles for all time. If the center of mass is at the origin and the triangle is symmetric about the second coordinate axis, the position vectors of the bodies may be written

$$q_1 = (-\frac{x}{2}, -\frac{m_3 y}{2+m_3}) \qquad q_2 = (\frac{x}{2}, -\frac{m_3 y}{2+m_3}) \qquad q_3 = (0, \frac{2y}{2+m_3})$$

where x, y are the Jacobi coordinates with

$$q_2 - q_1 = (x, 0) \qquad q_3 - \frac{1}{2}(q_1 + q_2) = (0, y) \qquad x \geq 0.$$

The isosceles three body problem is a Lagrangian system of two degrees of freedom with Lagrangian

$$\begin{aligned} L(x, y, \dot{x}, \dot{y}) &= \frac{1}{4}\dot{x}^2 + \frac{\mu}{2}\dot{y}^2 + U(x, y) \\ U(x, y) &= \frac{1}{r_{12}} + \frac{m_3}{r_{13}} + \frac{m_3}{r_{23}} \\ r_{12} &= x \quad r_{13} = r_{23} = \sqrt{\frac{x^2}{4} + y^2} \end{aligned} \tag{1}$$

where $\mu = \frac{2m_3}{2+m_3}$.

Date: July 18, 2019.

Let $r^2 = \frac{1}{2}|x_1|^2 + \mu|x_2|^2$ be the moment of inertia and define blown-up coordinates (r, θ) by

$$x = \sqrt{2} r \cos\theta \qquad y = \frac{1}{\sqrt{\mu}} r \sin\theta.$$

The polar variable r gives the size of the triangle formed by the three bodies while θ represent the shape. The Lagrangian becomes

$$\begin{aligned} L(r, \theta, \dot r, \dot\theta) &= \frac{1}{2}(\dot r^2 + r^2\dot\theta^2) + \frac{1}{r} V(\theta) \\ V(\theta) &= \frac{1}{r_{12}} + \frac{2m_3}{r_{13}} \\ r_{12} = \sqrt{2}\cos\theta &\quad r_{13} = \sqrt{\frac{\cos^2\theta}{2} + \frac{\sin^2\theta}{\mu}} \end{aligned} \tag{2}$$

Since $x \geq 0$, we have $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. Solutions of the Euler-Lagrange equations preserve the energy

$$\frac{1}{2}(\dot r^2 + r^2\dot\theta^2) - \frac{1}{r} V(\theta) = h.$$

The *shape potential* $V(\theta)$ for the isosceles problem is shown in Figure 1 for the case $m_3 = 1$. Note that $\theta = \pm\frac{\pi}{2}$ corresponds to $x = r_{12} = 0$, a double collision of masses m_1, m_2. The collision occurs on the second coordinate axis in the plane with mass m_3 above the colliding masses for $\theta = \frac{\pi}{2}$ and below them for $\theta = -\frac{\pi}{2}$. The three critical points correspond to three central configurations. The local maximum at $\theta = 0$ represents the Euler central configuration which is a collinear shape with m_3 at the midpoint of m_1, m_2. The two local minima represent the Lagrangian equilateral triangle shapes with m_3 above or below the other two masses. These general features of the shape potential do not depend on m_3 but the value of θ at the equilateral shapes varies.

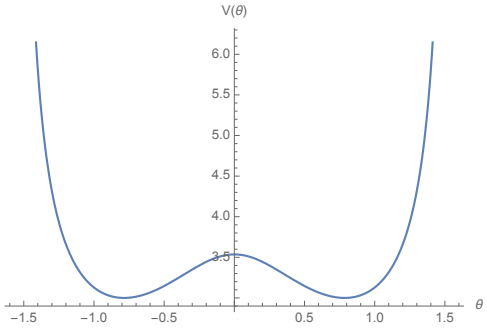

FIGURE 1. Shape potential for the equal mass isosceles three-body problem ($m_3 = 1$). $V(\theta)$ has singularities corresponding to double collisions at $\pm\frac{\pi}{2}$ and critical points at the three central configurations.

The use of polar variables is one step in the McGehee blowup procedure [8, 9]. In addition, we will introduce the McGehee timescale τ with

$$\frac{d\tau}{dt} = r^{\frac{3}{2}}$$

and define new velocity variables

$$v = \sqrt{r}\,\dot{r} \qquad w = r^{\frac{3}{2}}\dot{\theta}.$$

Then the blown-up Euler Lagrange differential equations are

$$\begin{aligned} r' &= v\,r \\ v' &= \frac{1}{2}v^2 + w^2 - V(\theta) \\ \theta' &= w \\ w' &= -\frac{1}{2}vw + V'(\theta). \end{aligned} \tag{3}$$

The energy equation can be written

$$\frac{1}{2}(v^2 + w^2) - V(\theta) = rh. \tag{4}$$

2. Geodesics of the Jacobi-Maupertuis Metric

In this section we give a brief discussion of the relation between solutions of the Euler-Lagrange equations and geodesics of the Jacobi-Maupertuis metric. Consider a Lagrangian system with Lagrangian of the form

$$L(q, v) = \frac{1}{2}\|v\|^2 + U(q)$$

with configuration $q \in \mathcal{U} \subset \mathbb{R}^n$ and velocity $v \in \mathbb{R}^n$. The first term is the kinetic energy and the second is minus the potential energy. Assume for simplicity that the kinetic energy takes the form

$$\frac{1}{2}\|v\|^2 = \frac{1}{2}v^T M v$$

where M is a positive definite symmetric matrix. Then solutions of the Euler-Lagrange equations with energy h satisfy

$$\begin{aligned} M\dot{q} &= p \\ \dot{p} &= \nabla U(q) \\ \frac{1}{2}\|\dot{q}\|^2 - U(q) &= h. \end{aligned} \tag{5}$$

The corresponding Jacobi-Maupertuis (JM) metric is a Riemannian metric on the *Hill's region*

$$\mathcal{H}(h) = \{q \in \mathcal{U} : U(q) + h \geq 0\}$$

given by

$$\mathbf{g}(v, v) = 2(U(q) + h)\|v\|^2.$$

If $\gamma(s)$ is a piecewise smooth curve in $\mathcal{H}(h)$ defined on a domain $\mathcal{D} \subset \mathbb{R}$ and if $[a, b] \subset \mathcal{D}$, then the arclength of γ between $\gamma(a), \gamma(b)$ is

$$l(\gamma, [a, b]) = \int_a^b \sqrt{\mathbf{g}(\gamma'(s), \gamma'(s))}\, ds.$$

If the interval $[a, b]$ is clear from context, this will be abbreviated to $l(\gamma)$. The arclength is invariant under reparametrization of the curve segment between $\gamma(a), \gamma(b)$ but it usual to use the unit speed parametrization which satisfies

$$\mathbf{g}(\gamma'(s), \gamma'(s)) = 1 \qquad l(\gamma, [a, b]) = b - a.$$

Unit speed geodesics can be characterized as curves swept out by free motions in $\mathcal{H}$ where a free motion is defined as a solution of the Euler-Lagrange equations for the Lagrangian

$$\tilde{L}(q,v) = \frac{1}{2}\mathbf{g}(v,v) = (U(q)+h)\|v\|^2$$

consisting only of a kinetic energy term based on the JM metric. Using the unit speed condition, it is easy to check that the Euler-Lagrange equations can be written

$$\begin{aligned} Mq' &= \frac{p}{2(U(q)+h)} \\ p' &= \frac{\nabla U(q)}{2(U(q)+h)} \\ 2(U(q)+h)\|q'(s)\|^2 &= 1 \end{aligned} \tag{6}$$

where $p = \tilde{L}_v = 2(U(q)+h)Mv$. Comparison of equations (5) and (6) shows that the solutions differ only by a change of timescale. In other words, after reparametrization, the curves $q(t)$ arising from solutions of the Lagrangian $L(q,v)$ with energy h become unit speed geodesics of the corresponding JM metric.

For the isosceles problem in Cartesian coordinates (1) the JM metric can be written

$$\mathbf{g} = 2(U(x,y)+h)(\frac{1}{2}dx^2 + \mu dy^2)$$

and the Hill's region is $\mathcal{H}(h) = \{(x,y) : x > 0, U(x,y)+h \geq 0\}$. For the blown-up problem in polar coordinates (2) we have the metric

$$\mathbf{g} = \frac{2}{r}(V(\theta)+rh)(dr^2 + r^2 d\theta^2)$$

and the Hill's region is $\mathcal{H}(h) = \{(\theta,r) : r > 0, -\frac{\pi}{2} < \theta < \frac{\pi}{2}, V(\theta)+rh \geq 0\}$. These are shown in Figure 2 for negative energy, $h < 0$. In this case we have a curve where the conformal factor $2(U(x,y)+h)$ or $2(V(\theta)/r+h)$ vanishes. The length of this *Hill boundary curve* is zero and this produces problems for the corresponding theory of geodesics.

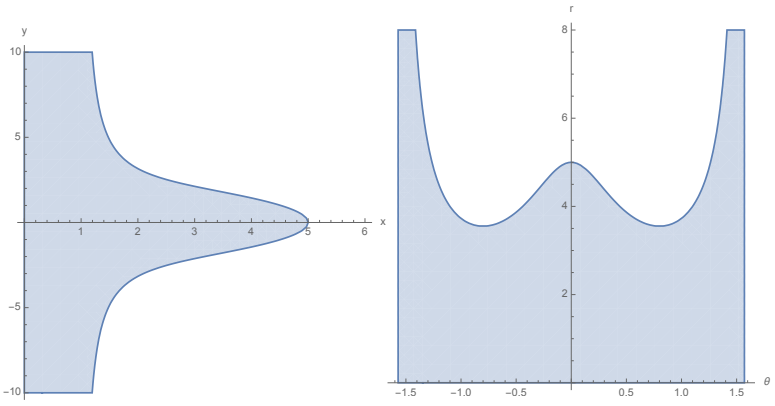

FIGURE 2. Hill's regions for the negative energy isosceles three-body problem in Cartesian and polar coordinates.

When $h \geq 0$ there is no Hill boundary curve although there are still singularities corresponding to double and triple collision where the conformal factor is infinite. It turns out that for the theory of geodesics, these singularities are less problematic than the Hill boundary curve. Therefore, from now on, will will focus on the closures of the Hill's regions for $h \geq 0$

$$\overline{\mathcal{H}}(h) = \{(x,y) : x \geq 0\} \qquad \tilde{\mathcal{H}}(h) = \{(\theta, r) : r \geq 0, -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}\} \qquad h \geq 0.$$

Clearly the Cartesian version is just the closed right half plane and the polar one is a closed vertical strip. Most of the theory is devoted to the Cartesian version but the polar one will be useful for some of the proofs and for plotting pictures of the geodesics later on.

From a purely geometric point of view, it's interesting to compute the Gaussian curvature of the two-dimensional Riemannian manifold $(\overline{\mathcal{H}}(h), \mathbf{g})$. Using standard formulas we find

$$K = \frac{V'(\theta)^2 - V(\theta)V''(\theta) + hr(V(\theta) + V''(\theta))}{2r(hr + V(\theta))^2}. \tag{7}$$

Later on, the sign of K will play an important role.

3. Jacobi-Maupertuis Distance

The distance between two points in a Riemannian manifold is usually defined as the infimum of the arclengths of all piecewise smooth curves connecting the points. We will see that in spite of the singularities of the JM metric at collisions, there is a well-defined Jacobi-Maupertuis distance between any two points of the closed Hill's region $\overline{\mathcal{H}}(h)$, $h \geq 0$.

The arclength of a piecewise smooth curve $\gamma(s) = (x(s), y(s))$ is given by

$$l(\gamma, [a,b]) = \int_a^b \sqrt{(U(x(s),y(s)) + h)(x'(s)^2 + 2\mu y'(s)^2)}\, ds. \tag{8}$$

The integrand is a Lebesgue integrable function even on the closed Hill's region, but the value of the arclength may be infinite. One way to deal with the fact that the potential is infinite on the boundary of $\overline{\mathcal{H}}(h)$ is to consider sequence of truncated potentials $U_N(x,y) = \phi_N(U(x,y))$ where $\phi_N : [0,\infty) \to [0, N+1]$ is some smooth, bounded cutoff function satisfying $\phi_N(u) = u$ for $0 \leq u \leq N$ and $\phi(u) \leq N+1$ for all u. Then for each $N > 0$

$$\mathbf{g}_N(v,v) = 2(U_N(x,y) + h)\|v\|^2$$

is a smooth Riemannian metric on $\overline{\mathcal{H}}(h)$. The arclengths $l_N(\gamma, [a,b])$ are all finite and the monotone convergence theorem shows that

$$l(\gamma, [a,b]) = \lim_{N \to \infty} l_N(\gamma, [a,b]).$$

The *Jacobi-Maupertuis distance* is defined as

$$d_{JM}(p,q) = \inf l(\gamma, [a,b])$$

where the infimum is over the set of all piecewise smooth curves in $\overline{\mathcal{H}}(h)$ with $\gamma(a) = p$ and $\gamma(b) = q$. The main result of this section is

Theorem 1. *For energies $h \geq 0$, d_{JM} is a well-defined distance function on $\overline{\mathcal{H}}(h)$ and $(\overline{\mathcal{H}}(h), d_{JM})$ is a complete metric space. Moreover, the topology induced on $\overline{\mathcal{H}}(h)$ by the JM metric agrees with the usual one.*

The proof will be done through a series of propositions.

Proposition 1. *For $p, q \in \overline{\mathcal{H}}(h)$, the JM distance satisfies $0 \le d_{JM}(p,q) < \infty$ with $d_{JM}(p,q) > 0$ if $p \ne q$,*

Proof. In addition to (8) is will be convenient to use the arclength formula in polar coordinates

$$l(\gamma, [a,b]) = \int_a^b \sqrt{2(V(\theta(s)) + r(s)h)(r'(s)^2/r(s) + r(s)\theta'(s)^2)}\, ds. \tag{9}$$

However, this is just a convenience and we still regard the underlying space $\overline{\mathcal{H}}(h)$ as a closed Cartesian halfplane. Using either formula, it is clear that $l(\gamma,[a,b]) \ge 0$ and so $d_{JM}(p,q) \ge 0$. To prove that the distance is finite, it suffices to find a curve $\gamma(s)$ with finite length connecting any two given points $p, q \in \overline{\mathcal{H}}(h)$. In fact, using polar coordinates, any two points can be connected by a sequence of horizontal and vertical line segments not lying entirely in the boundary. So it suffices to show that all such curves have finite length.

A vertical line segment not in the boundary can be parametrized with $r(s) = s, \theta(s) = \theta_0$ where $r_0 \le s \le r_1$ and $-\frac{\pi}{2} < \theta_0 < \frac{\pi}{2}$. Then the length satisfies

$$\begin{aligned} l(\gamma, [r_0, r_1]) &= \int_{r_0}^{r_1} \sqrt{\frac{2(V(\theta_0) + sh)}{s}}\, ds \le \int_{r_0}^{r_1} \sqrt{\frac{2(V(\theta_0) + r_1 h)}{s}}\, ds \\ &= 2\sqrt{2(V(\theta_0) + hr_1)}(\sqrt{r_1} - \sqrt{r_0}) < \infty \end{aligned} \tag{10}$$

even if one endpoint is at triple collision, $r_0 = 0$.

Similarly, a horizontal polar line segment not in the boundary can be parametrized with $r(s) = r_0 > 0, \theta(s) = s$, $-\frac{\pi}{2} \le \theta_0 \le s \le \theta_1 \le \frac{\pi}{2}$. The length of such a curve satisfies

$$l(\gamma, [\theta_0, \theta_1]) \le 2l(\gamma, [0, \frac{\pi}{2}]) = 2\int_0^{\frac{\pi}{2}} \sqrt{2r_0(V(\theta(s)) + r_0 h)}\, ds. \tag{11}$$

Along this curve, we have

$$V(\theta(s)) = \frac{1}{\sqrt{2}\cos s} + \frac{2m_3}{\sqrt{\frac{\cos^2 s}{2} + \frac{\sin^2 s}{\mu}}}.$$

The second term is bounded and the first has a singularity of order $O(1/(\frac{\pi}{2} - s))$ at $s = \frac{\pi}{2}$. It follows that the singularity in the length integral is of order $O(1/\sqrt{\frac{\pi}{2} - s})$ and the length is finite. For later use we also note that for small $\delta > 0$

$$\int_{\frac{\pi}{2}-\delta}^{\frac{\pi}{2}} \sqrt{2r_0(V(s) + r_0 h)}\, ds \le k(r_0)\sqrt{\delta} \tag{12}$$

for some constant $k(r_0)$.

If remains to show that $d_{JM}(p,q) > 0$ when $p \ne q$. This follows easily from the fact that the integrand in the length integral is strictly positive. Namely given p, q with $p \ne q$ consider a Euclidean ball of radius r around p in $\overline{\mathcal{H}}(h)$,

$$B(r,p) = \{p' \in \overline{\mathcal{H}}(h) : |p - p'| < r\}.$$

If r is sufficiently small, then the ball will not contain q. Moreover, since the closed ball is compact, there will be a positive constant k such that

$$k^2|v|^2 \le \mathbf{g}(v,v) \tag{13}$$

This allows comparison of the JM and Euclidean arclengths. Now every curve $\gamma(s)$ from p to q must go from p to the boundary of the ball and the JM length of the curve will be bounded below by kr. Hence $d_{JM}(p,q) \geq kr > 0$. QED

Proposition 1 shows that d_{JM} satisfies the positivity axiom for a metric space. The symmetry property $d_{JM}(p,q) = d_{JM}(q,p)$ is clear and the triangle inequality follows from the fact that the concatenation of piecewise smooth paths is another piecewise smooth path. Hence $(\overline{\mathcal{H}}(h), d_{JM})$ is a metric space.

Proposition 2. *The topology induced on $\overline{\mathcal{H}}(h)$ by the Jacobi-Maupertuis metric d_{JM} agrees with the usual topology as a closed half-space of $\mathbb{R}^2$.*

Proof. Let $B(r,p) = \{p' \in \overline{\mathcal{H}}(h) : |p - p'| \leq r\}$ denote a Euclidean ball of radius r and let $C(r,p) = \{p' \in \overline{\mathcal{H}}(h) : d_{JM}(p,p') \leq r\}$ be a ball with respect to the JM metric. Let $p \in \overline{\mathcal{H}}(h)$ and $B(r,p)$ be any Euclidean ball around p. We will show that there exists $r_1 > 0$ such that $C(r_1,p) \subset B(r,p)$. Similarly, given any JM ball $C(r,p)$ there is $r_1 > 0$ such that $B(r_1,p) \subset C(r,p)$.

The first part follows from the lower bound (13). Let $\gamma(s)$ be any curve segment from p to a point on the boundary of $B(r,p)$. Then the arclength of γ satisfies $l(\gamma) \geq kr$. It follows that for any point $q \notin B(r,p)$ we have $d_{JM}(p,q) \geq kr$ and hence $C(kr,p) \subset B(r,p)$.

For the second part, we are given a JM ball $C(r,p)$. First suppose $p \in \mathcal{H}(h)$ is a nonsingular points. By the first part of the proof, and taking r smaller if necessary, we may assume that $C(r,p) \subset B(r_2,p)$ and that $B(r_2,p) \subset \mathcal{H}(h)$. Then there will be a positive constant K such that

$$\mathbf{g}(v,v) \leq K^2|v|^2. \tag{14}$$

Let $r_1 = \min(r_2, r/K)$. For $q \in B(r_1,p)$, let γ be the line segment from p to q. Then

$$l(\gamma) < Kr_1 \leq r$$

and if follows that $q \in C(r,p)$ as required.

Next suppose p is the triple collision point, represented in polar coordinates by $\{r = 0\}$. Let $q = (r_0, \theta_0) \in B(r_1,p)$ where $0 < r_0 < r_1$. r_1 will be chosen later but for now we may assume $r_1 \leq 1$. Now q can be connected to p by the concatenation $\gamma = \gamma_1 + \gamma_2$ where γ_1 is the horizontal polar segment from q to the point with $r = r_0, \theta = 0$ and γ_2 is a vertical polar segment from the point $r = r_0, \theta = 0$ to the triple collision $r = 0, \theta = 0$. From (11) and our preliminary assumption $r_1 \leq 1$ we have a uniform estimate

$$l(\gamma_1) \leq k_1\sqrt{r_1} \qquad k_1 = 2\int_0^{\frac{\pi}{2}} \sqrt{2(V(s)+h)}\, ds.$$

On the other hand (10) shows

$$l(\gamma_2) \leq k_2\sqrt{r_1} \qquad k_2 = 2\sqrt{2(V(0)+h)}.$$

Hence $l(\gamma) \leq (k_1 + k_2)\sqrt{r_1}$. If we choose $r_1 < (\frac{r}{k_1+k_2})^2$, then we have $B(r_1,p) \subset C(r,p)$ as required.

Finally, suppose p is a double collision point, say $p = (r_0, \frac{\pi}{2})$ in polar coordinates. Given $C(r,p)$, it suffices to find a polar rectangle

$$R(\delta) = [r_0 - \delta, r_0 + \delta] \times [\frac{\pi}{2} - \delta, \frac{\pi}{2}] \subset C(r,p).$$

If $q = (r_1, \theta_1) \in R(\delta)$ then p can be connected to q by the concatenation $\gamma = \gamma_1 + \gamma_2 + \gamma_3$ where γ_1 is the horizontal polar segment from p to the point with polar coordinates $(r_0, \frac{\pi}{2} - \delta)$ and γ_2 is a vertical polar segment from $(r_0, \frac{\pi}{2} - \delta)$ to $(r_1, \frac{\pi}{2} - \delta)$ and γ_3 is a horizontal segment from $(r_1, \frac{\pi}{2} - \delta)$ to q. Using (12) we find

$$l(\gamma_1) \le k(r_0)\sqrt{\delta} \qquad l(\gamma_3) \le k(r_0)\sqrt{\delta}.$$

Using the fact that $V(\frac{\pi}{2} - \delta) = O(1/\sqrt{\delta})$, (10) gives a constant k_2 such that

$$l(\gamma_2) \le \frac{k_2}{\sqrt{\delta}}(\sqrt{r_0 + \delta} - \sqrt{r_0}) = O(\sqrt{\delta}).$$

It follows that if we choose δ sufficiently small, we will have $l(\gamma) < r$ and $R(\delta) \subset C(r, p)$ QED

To complete the proof of Theorem 1 we need one more proposition.

Proposition 3. *A subset $S \subset \overline{\mathcal{H}}(h)$ is bounded with respect to JM distance if and only if it is bounded with respect to the Euclidean distance.*

Proof. It suffices to consider balls centered at the triple collision point p. Consider the Euclidean ball $B(R, p)$ and let $q = (r_0, \theta_0)$, $0 < r_0 \le R$. Let γ_1 be the vertical polar segment from $r = \theta = 0$ to $r = r_0, \theta = 0$ and let γ_2 be the horizontal polar segment from $r = r_0, \theta = 0$ to q. Then (10) gives

$$l(\gamma_1) \le 2\sqrt{2(V(0) + hR)}\sqrt{R}$$

and (11) gives

$$l(\gamma_2) \le 2\int_0^{\frac{\pi}{2}} \sqrt{2R(V(s) + Rh)}\, ds.$$

Together, these give a bound on $d_{JM}(p, q)$ in $B(R, p)$.

Next consider the point q with polar coordinates R_1, θ. For any curve γ from p to q we have

$$l(\gamma) \ge \int_0^{R_1} \sqrt{\frac{2V_{min}}{s}}\, ds = 2\sqrt{2V_{min}}\sqrt{R_1}$$

where V_{min} is the minimal value of $V(\theta)$. For $q \in C(R, p)$, the JM ball centered at p, we have the Euclidean bound

$$R_1 \le \frac{R^2}{8V_{min}}.$$

QED

Proof of Theorem 1. It remains to show that the metric space $(\overline{\mathcal{H}}(h), d_{JM})$ is complete. Suppose p_n is a Cauchy sequence in $(\overline{\mathcal{H}}(h), d_{JM})$. In particular, it is bounded with respect to d_{JM}. By Proposition 3 p_n lies in some closed ball $B(R, p)$ around the triple collision point. This ball is compact with respect to the usual topology so by Proposition 2, it is also compact with respect to the topology induced by d_{JM}. Since compactness implies sequential compactness, p_n has a convergent subsequence. Since p_n is a Cauchy sequence, the sequence itself must be convergent. QED

4. Length Structure and Minimal Geodesics

We have used the arclength of piecewise continuous curves to define the JM distance function d_{JM} which makes the closed Hill's region into a metric space. In any metric space, one can define the length of arbitrary continuous curves. If $\gamma(s)$ is a continuous curve in $\overline{\mathcal{H}}(h)$ defined on an open domain $\mathcal{D} \subset \mathbb{R}$ and if $[a,b] \subset \mathcal{D}$, then the length of γ between $\gamma(a), \gamma(b)$ is

$$l(\gamma, [a,b]) = \sup \sum_{i=0}^{n-1} d_{JM}(\gamma(t_i), \gamma(t_{i+1})) \tag{15}$$

where the supremum is over all finite partitions $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ of $[a,b]$. As before we use the notation $l(\gamma)$ if the interval is understood.

We will need some ideas from the theory of *length structures*. A good reference is [2], especially Chapter 2. A length structure on a topological space is a class of curves and a length function obeying certain natural axioms. For example, arclength (8) on the class of piecewise smooth curves with $l(\gamma) < \infty$ gives a length structure on $\overline{\mathcal{H}}(h)$. On the other hand, if the underlying space is a metric space, there is a length structure on the class of continuous curves given by (15). Starting from the length structure on piecewise smooth curves we defined the JM metric and then using the metric we defined another length structure on the continuous curves. These two length structures agree on piecewise smooth curves. To see this we invoke [2, Theorem 2.4.3] which shows that it suffices to show that the piecewise smooth length structure is lower semi-continuous in the following sense: if γ_n is a sequence of piecewise smooth curves with finite length which converges pointwise to another such curve then $\liminf l(\gamma_n) \geq l(\gamma)$. This is a known result for smooth Riemannian metrics [2, Exercise 5.1.1]. To see that it holds for the JM metric, we will use that fact that for piecewise smooth curves, $l(\gamma) = \lim_{N\to\infty} l_N(\gamma)$ where l_N is the length with respect to the truncated metric g_N from the previous section. It is no loss of generality to assume a priori that $l(\gamma_n)$ $le 2l(\gamma)../$

Proposition 4. *If γ_n is a sequence of piecewise smooth curves with finite length which converges pointwise to another such curve then* $\liminf l(\gamma_n) \geq l(\gamma)$.

Proof. To see this, let $F(q) = \sqrt{2(U(q)+h)}$ be the conformal factor in the JM length integral. Then write $l(\gamma_n) - l(\gamma) = I_1 + I_2$ where

$$I_1 = \int_a^b F(\gamma_n(t)) \left(\|\gamma_n(t)'\| - \|\gamma(t)'\| \right) dt$$
$$I_2 = \int_a^b \left(F(\gamma_n(t)) - F(\gamma(t)) \right) \|\gamma(t)'\| \, dt$$

We will show that as $n \to \infty$, $I_2 \to 0$ and $\liminf I_1 \geq 0$.

All of the curves have finite length $l(\gamma_n)$ and are therefore uniformly bounded in the Euclidean norm. Then there will be a positive lower bound $F(\gamma_n(t)) \geq k > 0$ valid for all n and all $t \in [a,b]$ and therefore $l(\gamma_n) \geq k l_{euc}(\gamma_n)$. It is no loss of generality to assume that $l(\gamma_n) \leq 2l(\gamma)$ for all n. Then we get an estimate

$$|\gamma_n(t_2) - \gamma_n(t_1)| \leq \int_{t_1}^{t_2} |\gamma_n'(t)| \, dt \leq \sqrt{t_2 - t_1}$$

QED

We defined the JM distance as

$$d_{JM}(p,q) = \inf l(\gamma)$$

where the infimum was over all piecewise smooth curves from p to q. Now that we have a definition of length for continuous curves it is natural to consider the infimum over all continuous curves. It turns out that the two infimums agree. This is another standard result [2, Proposition 2.3.12(2)].

We want to apply length space theory to study geodesics of the metric space $(\overline{\mathcal{H}(h)}, d_{JM})$ with its associated length structure. Given $p, q \in \overline{\mathcal{H}(h)}$ a continuous curve $\gamma(t), a \leq t \leq b$ is a *shortest curve* from p to q if $\gamma(a) = p, \gamma(b) = q$ and

$$l(\gamma, [a,b]) = d_{JM}(p,q).$$

A curve $\gamma(t)$ on a domain $\mathcal{D} \subset \mathbb{R}$ is a *geodesic* if it is locally a shortest curve, that is, for each $t_0 \in \mathcal{D}$ there is an open neighborhood J of t_0 in $\mathcal{D}$ such that for each $t_1, t_2 \in J$, $\gamma(t)$ is a shortest curve from $\gamma(t_1)$ to $\gamma(t_2)$. A shortest curve from p to q is must also be shortest between every pair of its points, so it is a geodesic. More generally, a continuous curve $\gamma(t)$ on a domain $\mathcal{D} \subset \mathbb{R}$ is a *minimal geodesic* if is a shortest curve between every pair of its points.

Studying minimal geodesics of the JM metric with energy h is equivalent to studying so-called free time minimizers for the usual action functional (for the Lagrangian $L + h$). There have been several recent studies from this point of view [4, 6, 12, 13].

As long as a geodesic $\gamma(t)$ does not encounter the singular set of the metric, then the usual local theory for Riemannian manifolds shows that $\gamma(t)$ is a smooth curve. If the curve is parametrized at unit speed then it is a free motion and satisfies the Euler-Lagrange equations (6).

The theory of length spaces can be applied to show the existence of minimal geodesics.

Theorem 2. *For every pair of distinct points $p, q \in \overline{\mathcal{H}(h)}$, there is at least one minimal geodesic from p to q.*

Proof. This a metric geometry variation on the Hopf-Rinow theorem of differential geometry for complete Riemannian manifolds. The Hopf-Rinow theorem does not apply verbatim to our situation because of the singular points on the boundary of $\overline{\mathcal{H}(h)}$.

By Theorem 1, $(\overline{\mathcal{H}(h)}, d_{JM})$ is a complete metric space and by Proposition 2 it is locally compact. Moreover, we have seen that every pair of points can be connected by a path of finite length. Then the theorem follows directly from [2, Theorem 2.5.23]. QED

This result applies even when p or q or both are singular points. In this case, the minimal geodesic necessarily has singular points at its endpoints. The rest of this section is devoted to showing that a minimal geodesic cannot have singular points in the interior of its domain. This is a variation on a celebrated result called Marchal's lemma which asserts that action minimizing solutions of the N-body problem in $\mathbb{R}^d$ cannot have interior singular points [7, 3]. A similar result for the Jacobi-Maupertuis metric of the planar three-body problem can be found in [11]. It seems, however, that the isosceles constraint renders the proof of Marchal's lemma inapplicable. In any case, we will provide more elementary proofs here.

First we show that there can be no interior binary collisions. Recall that in polar coordinates, the binary collisions occur at $r > 0, \theta = \pm\frac{\pi}{2}$.

Proposition 5. *Suppose $\gamma(t), a \le t \le b$ is a minimal geodesic and $t_0 \in (a, b)$. Then $\gamma(t_0)$ is not a binary collision point.*

Proof. Suppose $\gamma(t_0)$ has polar coordinates $r(t_0) = r_0 > 0$ and $\theta(t_0) = \frac{\pi}{2}$ for some $t_0 \in (a, b)$. We will show that $\gamma(t)$ is not a minimal geodesic. A similar discussion will take care of the case $\theta(t_0) = -\frac{\pi}{2}$.

Clearly a curve of finite length cannot remain at $\theta = \frac{\pi}{2}$ for t in a nontrivial interval. It follows that there is $\theta_0 < \frac{\pi}{2}$ and times $t_1 < t_0 < t_2$ such that $\theta(t_1) = \theta(t_2) = \theta_0$ while $\theta(t) \ge \theta_0$ for $t_1 \le t \le t_2$. In other words, the angle must increase from θ_0 to $\frac{\pi}{2}$ and then return to θ_0. Now the shape potential $V(\theta)$ is strictly increasing on the interval $[\theta_{lag}, \frac{\pi}{2})$ where θ_{lag} is the equilateral shape where $V(\theta)$ achieves its minimum (see Figure 1). If we choose the angle $\theta_0 \in [\theta_{lag}, \frac{\pi}{2})$ then we have $V(\theta(t)) \ge V(\theta_0)$ for $t_1 \le t \le t_2$. We will see that the corresponding segment of $\gamma(t)$ cannot be minimal.

Let $\hat{\gamma}(t)$, $t_1 \le t \le t_2$, be the curve with polar coordinates $r = r(t)$, $\theta(t) = \theta_0$. In other words, project this part of $\gamma(t)$ onto the vertical polar line $\theta = \theta_0$. First we will show that this projection operation does not increase the length of the curve segment. Then we show that $\hat{\gamma}(t)$ is not minimal to complete the proof.

By definition of the length of a continuous curve and of the JM metric, we can approximate the segment $\gamma(t)$, $t_1 \le t \le t_2$ by a piecewise smooth curve $\tilde{\gamma}(t)$ with

$$l(\tilde{\gamma}, [t_1, t_2]) \le l(\gamma, [t_1, t_2]) + \epsilon$$

where $\epsilon > 0$ is any given constant. Then the arclength of $\tilde{\gamma}$ satisfies

$$\begin{aligned} l(\tilde{\gamma}, [t_1, t_2]) &= \int_{t_1}^{t_2} \sqrt{2(V(\theta(t)) + r(t)h)(r'(t)^2/r(t) + r(t)\theta'(t)^2)}\, dt \\ &\ge \int_{t_1}^{t_2} \sqrt{2(V(\theta_0) + r(t)h)(r'(t)^2/r(t))}\, dt = l(\hat{\gamma}, [t_1, t_2]). \end{aligned}$$

Hence $l(\hat{\gamma}, [t_1, t_2]) \le l(\gamma, [t_1, t_2]) + \epsilon$ and since ϵ was arbitrary we get $l(\hat{\gamma}, [t_1, t_2]) \le l(\gamma, [t_1, t_2])$.

To see that $\hat{\gamma}$ is not minimal we will show that short minimal geodesics connecting nearby points of the vertical segment $\theta = t_0$ lie strictly in the region $\theta < \theta_0$. Therefore replacing a short segment of $\hat{\gamma}$ by such a minimal curve will lower the total length. Using the differential equations (3), consider a geodesic with $\theta(t_0) = \theta_0$ and $\theta'(t_0) = 0$, so $\gamma(t_0)$ is tangent to the vertical polar line $\theta = \theta_0$. The second derivative satisfies $\theta''(t_0) = w'(t_0) = V'(\theta_0) > 0$. In other words, geodesics move into $\{\theta > \theta_0\}$ in both forward and backward time. In fact, this reasoning applies to any $\theta_0 \in (\lambda_{lag}, \frac{\pi}{2})$. It follows that any geodesic with a vertical tangent in this region ($\theta(t_0) = \theta_0 \in (\lambda_{lag}, \frac{\pi}{2}), \theta'(t_0) = 0$) has the property that $\theta(t)$ increases monotonically from θ_0 to $\frac{\pi}{2}$ in both forward and backward time. Fix θ_0 and consider the geodesics with vertical tangents at $\theta_0 - \delta$ where $\delta > 0$. These must cross the vertical line $\theta = \theta_0$ twice at points p_1, p_2. By taking δ small enough we can ensure that the corresponding geodesic segment are arbitrarily short. By a standard fact from Riemannian theory, such short geodesic segments are minimizers and are strictly shorter than the vertical segment from p_1 to p_2, as required. QED

In Cartesian coordinates, the triple collision singularity occurs at the origin $(x, y) = (0, 0)$. It is well-known that solutions of the three body problem which approach the triple collision singularity have shapes which approach the set of central configurations. For the isosceles problem, this fact implies that the curve $r(t), \theta(t)$ in polar coordinates satisfies $r(t) \to 0$ and $\theta(t) \to \theta_c$ where θ_c is one of the three critical points of $V(\theta)$. This observation will be used to prove the next result.

Proposition 6. *Suppose $\gamma(t), a \le t \le b$ is a minimal geodesic and $t_0 \in (a, b)$. Then $\gamma(t_0)$ is not the triple collision point.*

Proof. First note that a minimal geodesic cannot have two triple collision times. More generally, such curve cannot pass through any point twice. Otherwise we can find a shorter curve by simply deleting the loop. So we may assume $r(t_0) = 0$ and $r(t_0) > 0$ for $t \ne t_0$. Consider an interval $t_1 \le t \le t_2$ around t_0 such that $r(t_1) = r(t_2) = r_0 > 0$ where $r_0 > 0$ is a sufficiently small constant. We will show that the segment $\gamma(t), t_1 \le t \le t_2$ is not a minimal geodesic.

Now the three central configuration shapes are $\theta_c = 0, \pm\theta_{lag}$ where $\pm\theta_{lag}$ are angles corresponding to the equilateral triangle shapes. In polar coordinates, the curve segment $\gamma(t) = (r(t), \theta(t)), t_1 \le t \le t_2$ approaches triple collision at $\theta(t_0) = \theta_1$ and leaves at $\theta(t_0) = \theta_2$ where $\theta_i \in \{0, \pm\theta_{lag}\}$. If r_0 sufficiently small we may assume $|\theta(t_i) - \theta_i| < \epsilon$ where $\epsilon > 0$ is any given constant.

Up to symmetry, there are three cases to consider: $\theta_1 = -\theta_{lag}, \theta_2 = \theta_{lag}$, $\theta_1 = 0, \theta_2 = \theta_{lag}$ and $\theta_1 = \theta_2$. In all cases, estimates will show that γ is not the shortest path from $\gamma(t_1)$ to $\gamma(t_2)$.

It is convenient to introduce a change of variables to carry out the estimates. Namely, define new radial and angular variables (s, ϕ) via

$$r = s^2 \qquad \theta = 2\phi.$$

This is the squaring map of the complex plane. The new Hill's region is given by the wedge with $-\frac{\pi}{4} < \phi < \frac{\pi}{4}$. The JM metric becomes

$$\mathbf{g} = 8(V(2\phi) + s^2 h)(ds^2 + s^2 d\phi^2)$$

which is nonsingular at $s = 0$ and conformal to the Euclidean metric.

In the first of our cases $\theta_1 = -\theta_{lag}, \theta_2 = \theta_{lag}$ we use the simple estimates

$$l(\gamma, [t_1, t_0]) \ge \sqrt{8V_{min}}\, s_0 \qquad l(\gamma, [t_0, t_2]) \ge \sqrt{8V_{min}}\, s_0$$

to get

$$l(\gamma[t_1, t_2]) \ge 2\sqrt{8V_{min}}\, s_0 = \sqrt{32V_{min}}\, s_0.$$

On the other hand, let γ_3 be the straight line path from $\gamma(t_1)$ to $\gamma(t_2)$ in the Cartesian wedge corresponding to the polar variables (s, ϕ). We will have $V(2\phi(t)) \le V(0)$ and $s \le s_0$ along the whole segment (see Figure 1). Hence

$$l(\gamma_3) \le \sqrt{8(V(0) + s_0 h)}\, d_{euc}(\gamma(t_1), \gamma(t_2))$$

where d_{euc} denotes the Euclidean length of the segment. As the wedge is given by $-\frac{\pi}{4} \le \phi \le \frac{\pi}{4}$, this length satisfies

$$d_{euc}(\gamma(t_1), \gamma(t_2)) \le 2s_0 \cos\frac{\pi}{4} = \sqrt{2} s_0$$

and we have

$$l(\gamma_3) \le \sqrt{16(V(0) + s_0 h)}\, s_0.$$

Using these estimates we see that $l(\gamma_3) < l(\gamma, [t_1, t_2])$ provided the ratio

$$\frac{(V(0) + s_0 h)}{V_{min}} < 2.$$

Now s_0 can be made arbitrarily small and some calculation shows that

$$V(0) = \frac{1 + 4m_3}{\sqrt{2}} \qquad V_{min} = \frac{(1 + 2m_3)^{\frac{3}{2}}}{(2 + m_3)^{\frac{1}{2}}}.$$

From this we find that the ratio $V(0)/V_{min}$ achieves its maximum of $9\sqrt{3}/7\sqrt{2} \simeq 1.188$ at $m_3 = \frac{5}{8}$. So the required estimate is easily satisfied if s_0 is sufficiently small.

The other cases $\theta_1 = 0, \theta_2 = \theta_{lag}$ and $\theta_1 = \theta_2$ can be handled using the same crude estimates, the only difference being that that $d_{euc}(\gamma(t_1), \gamma(t_2))$ is even smaller. Again, we find that γ is not minimal. QED

Theorem 2 shows that many minimal geodesics exist. Here is a simple example.

Example 1. Minimal homothetic orbits. Central configurations of the n-body problem give rise to simple, explicit solutions. Perhaps the simplest are the homothetic solutions for $h \geq 0$, where the size of the configuration varies over $r \in (0, \infty)$ while the shape remains constant. For the isosceles problem, there are solutions of (3) with constant $\theta(t) = \theta_c \in \{0, \pm\theta_{lag}\}$ and $w(t) = \theta'(t) = 0$. The variables (r, v) satisfy

$$r' = v\, r$$
$$v' = \frac{1}{2}v^2 - V(\theta_c) = rh.$$

Energy $h = 0$ is the simplest case. We have constant $v(\tau) = \pm v_c$ where $v_c = \sqrt{2V(\theta_c)}$, and then $r(\theta) = r_0 \exp(\pm v_c \theta)$. For $v = -v_c$ we have a solution coming from infinity as $\tau \to -\infty$ and tending to triple collision as $\tau \to \infty$ in the McGehee timescale. In the usual timescale, the triple collision happens at a finite time. For $v = +v_c$ the solutions emerges from triple collision and tends to infinity in forward time. For $h > 0$ the behavior is qualitatively the same, except the solution reaches infinity in finite time in the McGehee timescale. In the original timescale this happens as $t \to -\infty$ or $t \to \infty$.

For the isosceles problem there are six of these solutions depending on which of the three central configurations is used and whether the size is increasing or decreasing. In the blown-up phase space $(\theta, r) \in [-\frac{\pi}{2}, \frac{\pi}{2}] \times [0, \infty)$ these solutions trace out vertical lines at $\theta = 0, \pm\theta_{lag}$. It follows from the estimates above that the JM arclength of the parts of these geodesics near triple collision are finite but that of the parts approaching $r = \infty$ are infinite.

Using the fact that $V(\pm\theta_{lag}) = V_{min}$ is the minimum of the shape potential, it is easy to show that the homothetic geodesics γ_{lag} corresponding to the equilateral central configurations at $\theta = \pm\theta_{lag}$ are minimal. To see this, consider the arclength of any piecewise smooth curve $\gamma(t) = (r(t), \theta(t))$ connecting two points of one of these geodesics, say $p = (r(a), \theta(a)) = (r_0, \theta_{lag})$ and $q = (r(v), \theta(v)) = (r_1, \theta_{lag})$

where $a < b$ and $r_0 < r_1$. We have

$$\begin{aligned} l(\gamma) &= \int_a^b \sqrt{2(V(\theta(t)) + r(t)h)/r(t)}\sqrt{r'(t)^2 + r(t)^2\theta'(t)^2}\,dt \\ &\geq \int_a^b \sqrt{2(V_{min} + r(t)h)/r(t)}\sqrt{r'(t)^2}\,dt \geq \int_{r_0}^{r_1} \sqrt{2(V_{min} + rh)/r}\,dr. \end{aligned}$$

The last formula is exactly $l(\gamma_{lag})$, the arclength of the homothetic geodesic and it follows that

$$l(\gamma_{lag}) = d_{JM}(p, q).$$

Since $\theta = 0$ is a local maximum of the shape potential, one might expect that the collinear geodesic γ_{euler} would not be minimal. But it will turn out that this depends on choice of the mass parameter m_3. Surprisingly, γ_{euler} is minimal at least when $h = 0$ and m_3 is large (see Theorem 3 below).

Although we know that a minimal geodesic exists between any two points $p, q \in \overline{\mathcal{H}}(h)$ it is not easy to understand the behavior of these geodesics in general. In the next section we will consider the simplest case $h = 0$ where we can use the known behavior of orbits on the collision manifold to understand the geodesics.

5. Minimal geodesics for the zero energy isosceles problem

The three-body problem with energy $h = 0$ has a homogeneity property which other energy levels do not. This is most clearly seen using McGehee's blown-up coordinates in (3). When $r = 0$, the blown-up energy equation (4) is independent of the size variable r and it defines a two-dimensional *collision manifold* in (θ, v, w)-space (see Figure 3). When $h = 0$ the same equation holds independent of r and the differential equations for the remaining variables (θ, v, w) are also independent of r.

Suppose $\gamma(\tau) = (r(\tau), \theta(\tau), v(\tau), w(\tau))$ is any solution of (3). Then the projected curve $\tilde{\gamma}(\tau) = (\theta(\tau), v(\tau), w(\tau))$ is a solution in the collision manifold. On the other hand, every solution $\tilde{\gamma}$ in the collision manifold is the projection of infinitely many zero energy solutions where the size can be recovered from the differential equation $r' = vr$:

$$r(\tau) = r(0) \exp \int_0^\tau v(s)\,ds.$$

If $r(0) = 0$ we have an artificial solution at triple collision, but any $r(0) > 0$ leads to a real solution to the isosceles three-body problem. In that case the corresponding curve $(\theta(\tau), r(\tau))$ in $\overline{\mathcal{H}}(0) = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [0, \infty)$ parametrizes a geodesic of the zero energy JM metric

$$\mathbf{g} = U(x, y)(dx^2 + 2\mu dy^2) = \frac{2V(\theta)}{r}(dr^2 + r^2 d\theta^2).$$

Now the flow on the collision manifold has been well studied and much is known about it [14, 5, 10]. The goal of this section is to translate this knowledge into some theorems about the minimal geodesics whose existence was proven in the last section. We begin by describing some features of the collision manifold flow. In most of the previous work, solutions are continued beyond the inevitable double collisions using some sort of regularization technique. But the results of the previous section show that the continued geodesics would not be minimal. So for the purposes of this paper, we can work with the unregularized differential equations (3).

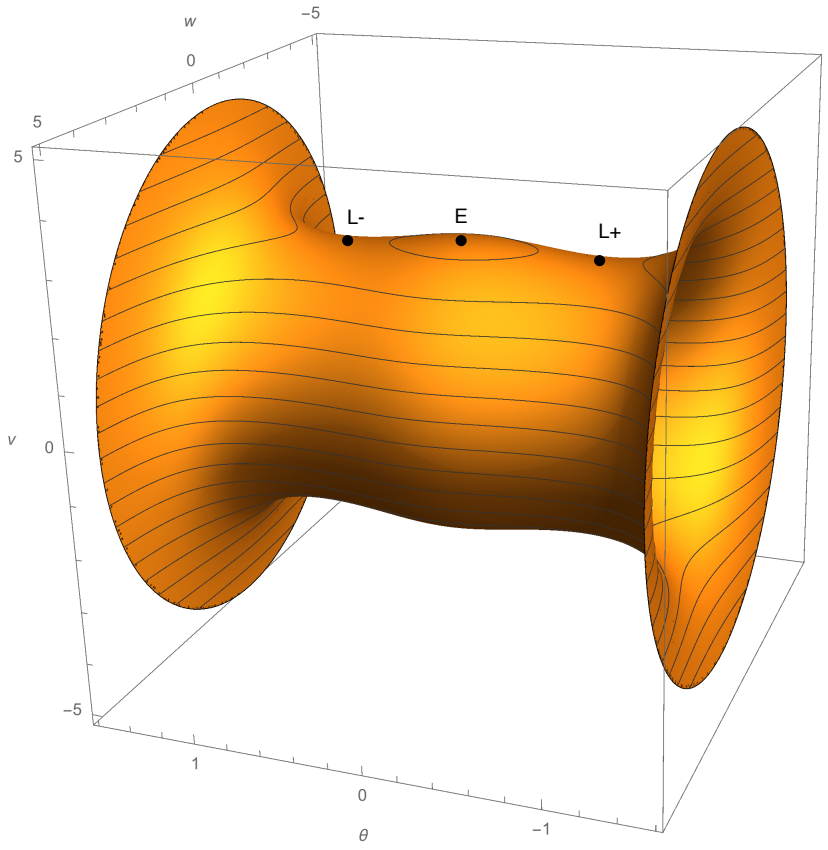

FIGURE 3. Unregularized collision manifold for the equal mass isosceles three-body problem ($m_3 = 1$). Level curves of the Lyapunov function v are shown.

Equations (3) have six equilibrium points given by $(r, \theta, v, w) = (0, \theta_c, v_c, 0)$ where $\theta_c \in \{0, \pm\theta_{lag}\}$ is one of the three central configurations and $v_c = \pm\sqrt{2V_c} = \pm\sqrt{2V(\theta_c)}$. The eigenvalues of the linearized equations corresponding to eigenvectors tangent to the energy manifold are

$$\lambda = v_c, \frac{-v_c + \sqrt{v_c^2 + 16V''(\theta_c)}}{4}, \frac{-v_c - \sqrt{v_c^2 + 16V''(\theta_c)}}{4}.$$

The eigenvector for $\lambda = v_c$ is the radial vector $(\delta r, \delta\theta, \delta v, \delta w) = (1, 0, 0, 0)$ and the other two are tangent to the collision manifold. For the equilibria with $v_c < 0$, the radial direction is attracting while for $v_c > 0$ it is repelling. The Euler equilibria

with $\theta_c = 0$ and $v > 0, v < 0$ will be labeled E, E^* respectively. The Lagrange equilibria with $\theta_c = \pm\theta_{lag}$ will be called $L_\pm, L_\pm^*$ with the star denoting $v < 0$.

The Lagrange central configurations are minima of V and $V''(\theta_c) > 0$. It follows that the eigenvalues of $L_\pm, L_\pm^*$ in the collision manifold are real and of opposite sign, so we have saddle points with one-dimensional stable and unstable manifolds. The Euler equilibria are local maxima of V and we have $V''(0) < 0$. Some computation shows that the quantity under the radical is negative for $m_3 < \frac{55}{4}$ and real for $m_3 \geq \frac{55}{4}$. In both cases, the two tangential eigenvalues are attracting at the $v_c > 0$ restpoints and repelling at the $v_c < 0$ ones. So E is either a stable spiral or a stable node while E^* is an unstable spiral or node.

Another important feature of the collision manifold flow is its gradient-like character. Using the energy equation, the equation for v' can be written

$$v' = \frac{1}{2}w^2 \geq 0$$

so $v(\tau)$ is nondecreasing along solutions. In fact $v(\tau)$ is strictly increasing except at the restpoints. To see this note that if $w(\tau_0) \neq 0$ we have $v'(\tau_0) > 0$. If $w(\tau_0) = 0$ but $\tilde{\gamma}(\tau_0)$ is not a restpoint then we have $w'(\tau_0) = V'(\theta(\tau_0)) \neq 0$. It follows that for time $\tau \neq \tau_0$ sufficiently close to τ_0 we have $w(\tau) \neq 0$ and so $v(\tau)$ is still strictly increasing. Level curves of v are shown in Figure 3 where the conclusion is that all solutions $\tilde{\gamma}(\tau)$ are moving steadily upward.

Next, note that solution with $w > 0$ (the front of the manifold in Figure 3) have $\theta'(\tau) > 0$ (so the move from left to right in the figure). Solutions on the back of the manifold, $w < 0$, move from right to left. Figure 4 shows stream lines of the flow on the front for $m_3 = 1$, projected to the (θ, v) plane. The flow on the back is the reflection of this one through the vertical coordinate axis. Many of the features evident in the figure actually hold for all $m_3 > 0$. Specifically, we will mainly use the some simple properties of the stable and unstable manifolds of the saddle points. Focussing on the front part of the manifold, denote the $w > 0$ branches of the stable and unstable manifolds of $L_\pm$ by $c_\pm^s, c_\pm^u$ and those of $L_\pm^*$ by $c_\pm^{s*}, c_\pm^{u*}$. Then using (θ, v) as coordinates on the front of the surface, we will use the following properties of the flow in the collision manifold:

- $c_\pm^s$ are graphs of the form $v = v_\pm^s(\theta)$ for $-\frac{\pi}{2} < \theta \leq \pm\theta_{lag}$ with $v_\pm^s(\theta)$ monotonically increasing and with $0 < v_+^s(\theta) < v_-^s(\theta)$.
- Similarly, $c_\pm^{u*}$ are curves of the form $v = v_\pm^{u*}(\theta)$ for $\pm\theta_{lag} \leq \theta < \frac{\pi}{2}$ with $v_\pm^{u*}(\theta)$ monotonically increasing and with $v_+^{u*}(\theta) < v_-^{u*}(\theta) < 0$.
- c_+^u is a curve of the form $v = v_+^u(\theta)$ for $\theta_{lag} \leq \theta < \frac{\pi}{2}$ with $v_+^u(\theta)$ monotonically increasing
- c_-^{s*} is a curve of the form $v = v_-^{s*}(\theta)$ for $-\frac{\pi}{2} < \theta \leq -\theta_{lag}$ with $v_+^{s*}(\theta)$ monotonically increasing

Together, these curves divide the front part of the collision manifold into seven disjoint regions (see Figure 4). In each region, the solutions have a similar qualitative behavior.

We want to relate the orbits on the collision manifold to the geodesics of the JM metric. The equilibrium points on the collision manifold correspond to the homothetic solutions and geodesics. Consider, for example, the equilibrium point L_+ on the collision manifold. The size variable is given by $r(\tau) = r(0)\exp(v_c\tau)$ where $v_c = \sqrt{2V_{min}}$ and $r(0) > 0$ is an arbitrary initial size. The corresponding geodesic is the vertical polar line $\theta(\tau) = \theta_{lag}$ parametrized in the sense of increasing

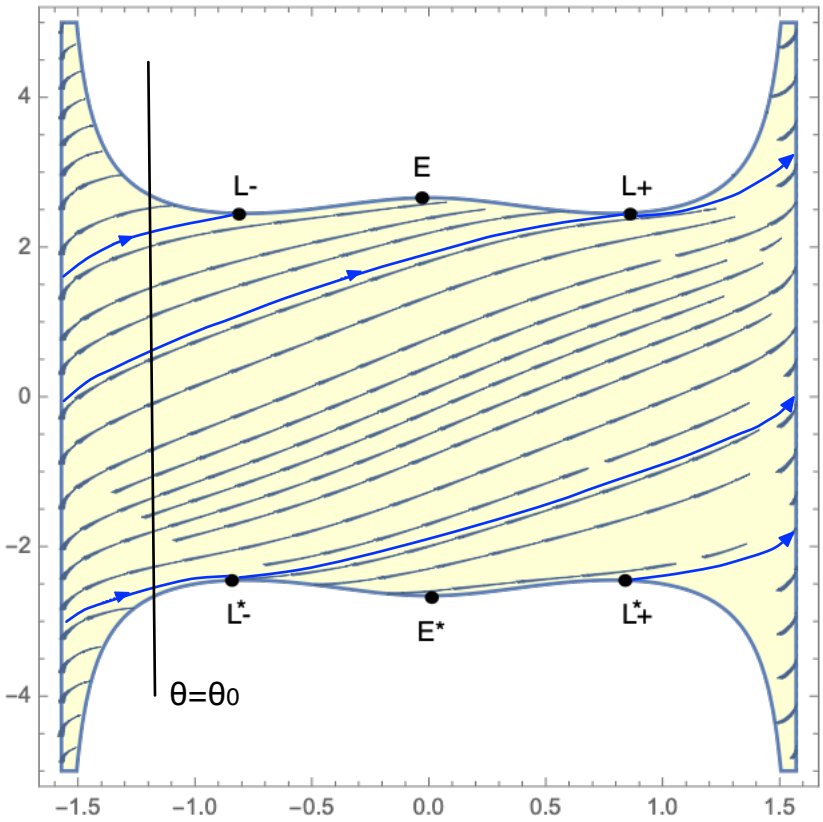

FIGURE 4. Projection of the flow on the front ($w > 0$) side of the collision manifold to the (θ, v) plane for $m_3 = 1$, showing the stable manifolds of some of the saddle points. A slice of the form $\theta = \theta_0$ is indicated.

r. The equilibrium L_+^* on the bottom half of the collision manifold leads to a parametrization of the same line in the sense of decreasing r. Similarly, the other equilibrium points produce parametrizations of the geodesics on the vertical lines $\theta = -\theta_{lag}$ and $\theta = 0$.

Next consider the geodesics beginning at an arbitrary point $p \in \overline{\mathcal{H}}(h)$, say the point with $r = r_0, \theta = \theta_0$. We know that there is a minimal geodesic connecting p to any other point $q \in \overline{\mathcal{H}}(h)$. If p is not a collision point, then there is a circle of possible initial velocities for a geodesic beginning at p, parametrized by the unit vectors in $\overline{\mathcal{H}}(h)$. Figure 5 shows numerically computed geodesics beginning at a

typical point and emerging in different directions for the case $m_3 = 1$. The observed behavior of these geodesics and their minimality properties can be understood with reference to the collision manifold flow of Figure 4.

The initial point p in Figure 5 is of the form $r_0 = 1, \theta_0 \in (-\frac{\pi}{2}, -\theta_{lag})$. Every geodesic emanating from p corresponds to an orbit $\tilde{\gamma} = (\theta(\tau), v(\tau), w(\tau))$ on the collision manifold beginning on the slice $\theta = \theta_0$. Clearly this slice is just a circle in the (v, w) plane which corresponds to the possible initial velocities for the geodesic. Geodesics with initial velocities with $w \geq 0$ correspond to solutions initially on the front of the manifold. In the projection to the (θ, v) plane these appear as all possible solutions beginning on the line segment $\theta = \theta_0$ and moving left to right (see Figure 4). Note that this line segment intersects three of the projected stable manifolds. From top to bottom they are c^s_-, c^+_s, c^{s*}_-. These orbits on the collision manifold correspond to the three red geodesics in Figure 5.

Consider, for example, the geodesic γ whose corresponding collision manifold orbit is in c^s_-. The corresponding orbit on the collision manifold has $\theta(\tau) \to -\theta_{lag}$ and $0 < v(\tau) \to \sqrt{2V_{min}}$ with both θ and v increasing monotonically from their initial values. Since $r' = vr$, $r(\tau)$ increases monotonically from r_0 to ∞. Hence the geodesic approaches $r = \infty$ with the shape converging to $-\theta_{lag}$ from the left, as indicated in the figure. It is not hard to see that γ is a minimal geodesic. Let q be any point on the geodesic. Since we know that there exists a minimal geodesic from p to q it suffices to show that γ is the *only* geodesic from p to q which has no collisions. Suppose the polar coordinates of q are r_1, θ_1. It suffices to consider the geodesics corresponding to orbits on the collision manifold which run from $\theta = \theta_0$ to $\theta = \theta_1$ without hitting $\theta = \pm\frac{\pi}{2}$. These must lie entirely on the front of the collision manifold. Note that all solutions on the front of the collision manifold can be written as graphs $v = v(\theta)$. The following lemma estimates the r coordinates of the corresponding geodesics.

Lemma 1. *Let $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ be two solutions with $w > 0$ and such that θ_i increases at least from θ_0 to θ_1. Parametrizing the solutions by $\theta \in [\theta_0, \theta_1]$, suppose that $r_1(\theta_0) = r_2(\theta_0)$ and $v_1(\theta_0) < v_2(\theta_0)$. Then $v_1(\theta) < v_2(\theta)$ and $r_1(\theta) < r_2(\theta)$ for $\theta \in (\theta_0, \theta_1]$.*

Proof. Since distinct orbits on the collision manifold do not intersect, we must have $v_1(\theta) < v_2(\theta)$, that is, γ_2 remains above γ_1. Since $r' = vr$ and $\theta' = w$ we have

$$\frac{dr_i}{d\theta} = \frac{v_i(\theta)}{w_i(\theta)} r_i.$$

Note that the slice of the collision manifold at fixed θ is a circle in the (w, v) plane. It follows when $w > 0$, points with larger v values also have larger values of v/w. Integrating the differential equation, we get $r_1(\theta) < r_2(\theta)$ for $\theta \in (\theta_0, \theta_1]$ as claimed QED

Returning to the proof of minimality, we see that for geodesics starting at p and running at least to $\theta = \theta_1$, the final radius is a monotonically increasing function of the initial slope. In particular, γ is the only geodesic from p to q and since a minimal geodesic exists, it must be minimal.

One can understand the other red geodesics in Figure 5 in a similar way. The geodesic whose corresponding collision orbit is in the stable branch c^s_+ is a minimal geodesic starting at p and converging to $r = \infty$ with asymptotic shape $+\theta_{lag}$ with

both θ and r increasing monotonically. On the other hand, the branch c_-^{s*} leads to a minimal geodesic converging to triple collision $r = 0$ with with asymptotic shape $-\theta_{lag}$ and with θ increasing and r decreasing monotonically.

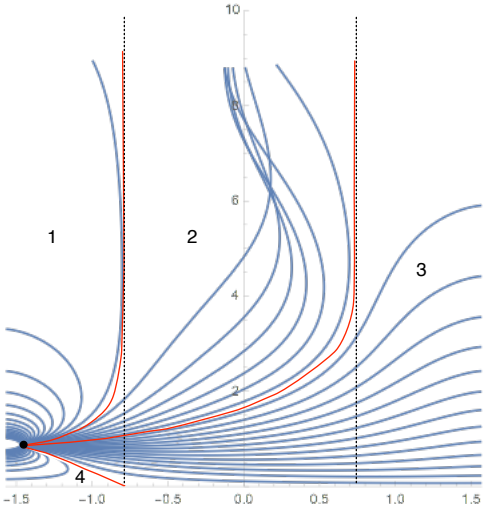

FIGURE 5. Polar plot of geodesics emanating from a typical non-singular point $m_3 = 1$.

The three red geodesics just discussed divide the set of all geodesics emanating from p with $w \geq 0$ into four disjoint regions, labeled 1-4 in Figure 5. We want to understand the minimal geodesics in each region. The easiest is region 3. The corresponding collision manifold orbits run all the way from $\theta = \theta_0$ to $\theta = \frac{\pi}{2}$. Lemma 1 shows that for a given $\theta \in (\theta_0, \frac{\pi}{2}]$, $r(\theta)$ is a monotonically increasing function of the initial slope. Note that these are the only geodesics which can reach points q with $\theta \geq \theta_{lag}$. Since we know that minimal geodesics from p to q exist, it follows that all of the geodesics in region 3 are minimal. Before discussing the other regions with $w \geq 0$, note that this same argument applies to geodesics at p with initial $w < 0$ whose corresponding collision manifold orbits are on the back of the collision manifold. By symmetry, studying these is equivalent to studying geodesics with $w > 0$ starting on the slice $\theta = -\theta_0 \in (\theta_{lag}, \frac{\pi}{2})$. These orbits move monotonically to double collision and the final size r is a monotonic function of the initial slope.

The geodesics in region 1 correspond to orbits on the front of the collision manifold with initial conditions above the stable branch c_-^s. These collision orbits proceed monotonically to the curve $w = 0$ and then move to the back of the collision manifold. It follows that the corresponding geodesics have θ and r increasing monotonically until a vertical tangent is reached, when they proceed with r still increasing and θ decreasing to the double collision at $\theta = -\frac{\pi}{2}$. We will show that

all of these geodesics are disjoint and also do not intersect the geodesics with initial $w < 0$. From this it will follow, as above, that they are all minimal.

The following lemma about the Gauss curvature gives the disjointness.

Lemma 2. *For $h = 0$ and $m_3 > 0$, the Gauss curvature is strictly negative for $-\frac{\pi}{2} < \theta \le \theta_{lag}$ and $\theta_{lag} \le \theta < \frac{\pi}{2}$. In particular, distinct geodesics remaining entirely in these regions must be disjoint.*

Proof. Equation (7) with $h = 0$ shows that the sign of the Gauss curvature is the same as the sign of $V'(\theta)^2 - V(\theta)V''(\theta)$. The first derivative of $V(\theta)$ can be written

$$V'(\theta) = \sin\theta \left(\frac{1}{\sqrt{2}\cos^2\theta} - \frac{2\cos\theta}{r_{13}^3} \right) \qquad r_{13}^2 = \frac{1}{2} + \frac{\sin^2\theta}{m_3}.$$

For $-\frac{\pi}{2} < \theta \le \theta_{lag}$ and $\theta_{lag} \le \theta < \frac{\pi}{2}$ we have

$$\frac{2\cos\theta}{r_{13}^3} < \frac{1}{\sqrt{2}\cos^2\theta} \qquad \sqrt{2}\cos\theta < r_{13}.$$

If we let $c = \cos\theta, s = \sin\theta$, we can write $V'(\theta)^2 - V(\theta)V''(\theta)$ as

$$s^2 \left(\frac{1}{\sqrt{2}c^2} - \frac{2c}{r_{13}^3} \right)^2 - \left(\frac{1}{\sqrt{2}c} + \frac{2m_3}{r_{13}} \right) \left(\frac{1}{\sqrt{2}c^3} - \frac{2c^2}{r_{13}^3} + \frac{s^2}{\sqrt{2}c^2} + \frac{2s^2}{r_{13}^3} + \frac{6c^2s^2}{m_3 r_{13}^5} \right) = A + s^2 B$$

where

$$A = -\left(\frac{1}{\sqrt{2}c} + \frac{2m_3}{r_{13}} \right) \left(\frac{1}{\sqrt{2}c^3} - \frac{2c^2}{r_{13}^3} \right)$$

$$B = \left(\frac{1}{\sqrt{2}c^2} - \frac{2c}{r_{13}^3} \right)^2 - \left(\frac{1}{\sqrt{2}c} + \frac{2m_3}{r_{13}} \right) \left(\frac{1}{\sqrt{2}c^2} + \frac{2}{r_{13}^3} + \frac{6c^2}{m_3 r_{13}^5} \right).$$

It turns out that after adding $\frac{s^2}{2c^4} = \frac{1-c^2}{2c^4}$ to A and subtracting $\frac{1}{2c^4}$ from B, both quantities are negative (of course this leaves $A + s^2 B$ unchanged). We have

$$(2c^2 r_{13}^3)(A + \frac{s^2}{2c^4}) = c r_{13}(2\sqrt{2}c^3 - r_{13}^3) + 2\sqrt{2}m_3(2\sqrt{2}c^5 - r_{13}^3) < 0$$

where the inequality follows from our assumption that $\sqrt{2}\cos t < r_{13}$. Also

$$2c^3 b_3 r_{13}^6 (B - \frac{1}{2c^4})$$

is a complicated polynomial in m_3, c, r_{13} with all nonpositive terms. It follows that the sign of the Gauss curvature is negative as required.

The claim about geodesics being disjoint is a corollary of the Gauss-Bonnet theorem. Namely, if two distinct geodesics starting at p intersect at another points q then they form a geodesic polygon P with two edges and two vertices. Let the interior angles be denoted $\alpha_p > 0$ and $\alpha_q > 0$. Then the Gauss-Bonnet theorem implies that

$$\alpha_p + \alpha_q = \int_P K \, dA < 0$$

a contradiction. QED

A similar argument applies to geodesics in region 4. These correspond to orbits on the front of the collision manifold with initial conditions below the stable branch c_-^{s*}. These collision orbits also proceed monotonically to the curve $w = 0$ and then move to the back of the collision manifold. It follows that the corresponding geodesics have θ increasing and and r decreasing monotonically until a downward vertical tangent is reached, when they proceed with r still decreasing and θ decreasing to the double collision at $\theta = -\frac{\pi}{2}$. The same argument as above shows that these geodesics do not intersect and are all minimal.

Geodesics in region 2 are more complicated. When $m_3 < \frac{55}{4}$ geodesics oscillate around the collinear homothetic geodesic intersecting it and one another multiple times. These will be minimal up to the first intersection point. The existence theorem 2 shows that there are minimal geodesics from region 2 connecting p to every point q in the vertical strip with $-\theta_{lag} \leq \theta \leq \theta_{lag}$. It is interesting to see from Figure 5 how it is possible to reach arbitrarily large r values by staying close to the Lagrange homothetic orbits at $\theta = \pm\theta_{lag}$. The details about the location of the intersections of geodesics from region 2 seem difficult to understand.

When $m_3 \geq \frac{55}{4}$, there are real eigenvalues at the collinear equilibrium points E, E^* and correspondingly, no oscillation of geodesics around the collinear, Euler homothetic geodesic. Figure 6 shows the geodesics emanating from the same point p when $m_3 = 20 > \frac{55}{4}$. This time, there is no oscillation around the Euler geodesic and no obvious reason for this geodesic to fail to be minimal. Indeed, we have

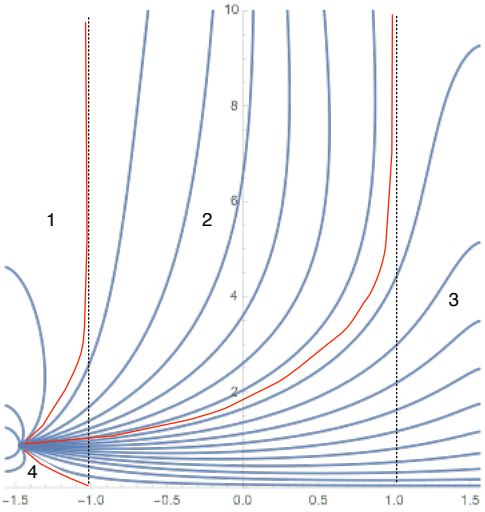

FIGURE 6. Polar plot of geodesics emanating from a typical non-singular point $m_3 = 20 > \frac{55}{4}$.

Theorem 3. *For $m_3 \geq \frac{55}{4}$ and energy $h = 0$, the geodesics corresponding to the Euler homothetic orbits are minimal in the isosceles problem.*

Proof. As above, the proof is based on uniqueness. We will see that if p, q are points along the line $\theta = 0$ then the Euler geodesic is the only geodesic running from p to q. Since a minimal geodesic exists, the Euler geodesic must be minimal. To prove uniqueness, note that any geodesic from p to q corresponds to a collision manifold

orbit beginning on the circle $\theta = 0$ and returning to this circle at a later time. Of course the two Euler restpoints E, E^* have this property, but we will see that no other such orbit exists.

Focussing on orbits on the front ($w > 0$) part of the collision manifold, it is clear from Figure 4 that the initial condition of a solution starting and returning to $\theta = 0$ would have to lie between the branches of the stable manifold saddle points L_-, L_+. Consider a solution starting on $\theta = 0$ between c^s_-, c^s_+. The hypothetical solution would have to move to the back of the collision manifold and then return to $\theta = 0$ with $w < 0$. Consider such a solution as it crosses the ray $w = 0, \theta > 0$ and moves into the region where $w < 0$. A glance at Figure 7 shows that such solutions apparently converge to the restpoint E while remaining in the region $\theta > 0$. Although they approach $\theta = 0$ asympotically, the never actually reach it. The rest of the proof amounts to providing a rigorous way to see that this actually occurs.

Let the initial condition of the solution in question be $w = 0, \theta = \theta_0$ where $0 < \theta_0 < \theta_{lag}$. The initial value $v = v_0$ is determined from the equation of the collision manifold. We will show that the solution enters a positively invariant trapping region entirely contained in $\{\theta > 0\}$. The trapping region is given by

$$\mathcal{T} = \{(\theta, w, v) : v \geq v_0, 0 < \theta \leq \theta_0, -k \sin\theta \leq w \leq 0\}$$

where $k > 0$ is an appropriate constant (see Figure 7). To see that the $\mathcal{T}$ is positively invariant it suffice to show that solutions on the boundary curves do not leave in forward time. Recall that v is a Lyapunov function and note that when $w = 0$ and $0 < \theta < \theta_{lag}$, $w' = V'(\theta) < 0$. So it only remains to check the boundary curve where $f(\theta, w, v) = w + k \sin\theta = 0$. The following lemma shows that $f' \geq 0$ when $f = 0$ and this completes the proof of Theorem 3. QED

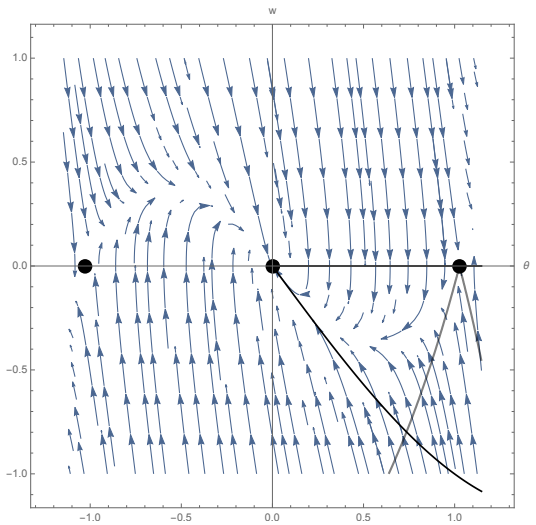

FIGURE 7. Solutions near the Euler equilibrium E when $m_3 = 20$ using coordinates (θ, w). The trapping region $\mathcal{T}$ is also shown. The lower left edge is the curve $f = w + k \sin\theta = 0$ and the lower right edge is $v = v_{lag}$.

Lemma 3. *Let $m_3 \geq \frac{55}{4}$ and let $f(\theta, w, v) = w + k\sin\theta$ where $-k$ is the weaker of the two attracting eigenvalues at the equilibrium point E. If (θ_0, w_0, v_0) is a point on the collision manifold with $0 < \theta_0 < \theta_{lag}$ and $f(\theta_0, w_0, v_0) = 0$, then $f'(\theta_0, w_0, v_0) \geq 0$.*

Proof. The differential equations (3) give

$$f' = w' + k\cos\theta\,\theta' = -\frac{1}{2}vw + V'(\theta) + k\cos\theta\, w.$$

If $f = 0$ then $w = -k\sin\theta$. Calculating $V'(\theta)$ leads to the formula

$$f'|_{f=0} = \sin\theta\, g(\theta) \qquad g(\theta) = \frac{1}{2}kv - k^2\cos\theta + \frac{1}{\sqrt{2}\cos^2\theta} - \frac{4\sqrt{2}\cos\theta}{d^3}$$

where

$$d = \sqrt{1 + 2\delta\sin^2\theta} \qquad \delta = \frac{1}{m_3}.$$

It suffices to show that $g(\theta) \geq 0$ when $0 \leq \theta \leq \frac{\pi}{2}$ and $0 \leq \delta \leq 4/55$.

The constant k is minus the weaker of the two attracting eigenvalues at E which works out to be

$$k = \frac{\sqrt{2V(0)}}{4}\left(1 - \sqrt{1 + 8\frac{V''(0)}{V(0)}}\right) = \frac{7\cdot 2^{\frac{1}{4}}\sqrt{\delta}}{r_1 + r_2}$$

where

$$r_1 = \sqrt{1 + \frac{\delta}{4}} \qquad r_2 = \sqrt{1 - \frac{55}{4}\delta}.$$

One can eliminate v from the first term of g using the energy relation $v^2 = 2V(\theta) - w^2 = 2V(\theta) - k^2\sin\theta^2$ which leads after some algebra to

$$\sqrt{\delta}v = \sqrt{\frac{4\sqrt{2}}{d} + \frac{\sqrt{2}\delta}{\cos\theta} - \frac{49\sqrt{2}\,\delta^2}{(r_1+r_2)^2}\sin^2\theta}.$$

Substituting this and the formula for k into g gives

$$g = \frac{7}{2^{\frac{3}{4}}(r_1+r_2)}\sqrt{\frac{4\sqrt{2}}{d} + \frac{\sqrt{2}\delta}{\cos\theta} - \frac{49\sqrt{2}\,\delta^2}{(r_1+r_2)^2}\sin^2\theta} - \frac{49\sqrt{2}\,\delta^2}{(r_1+r_2)^2}\cos^2\theta + \frac{1}{\sqrt{2}\cos^2\theta} - \frac{4\sqrt{2}\cos\theta}{d^3}.$$

The formula makes sense when $\delta = 0$ and gives

$$g_0 = \frac{7}{\sqrt{2}} + \frac{1}{\sqrt{2}\cos^2\theta} - 4\sqrt{2}\cos\theta = \frac{7}{\sqrt{2}}(1 - \cos\theta) + \frac{1}{\sqrt{2}\cos^2\theta}(1 - \cos^3\theta).$$

This expression shows that $g_0 \geq 0$. To complete the proof, it will be shown that $g \geq g_0$ for $0 \leq \delta \leq 4/55$.

The difference $g - g_0$ can be written $g - g_0 = h(\theta) + 4\sqrt{2}\cos\theta(1 - d^{-3})$ where

$$h(\theta) = \frac{7}{2^{\frac{3}{4}}(r_1+r_2)}\sqrt{\frac{4\sqrt{2}}{d} + \frac{\sqrt{2}\delta}{\cos\theta} - \frac{49\sqrt{2}\,\delta^2}{(r_1+r_2)^2}\sin^2\theta} - \frac{7}{\sqrt{2}} - \frac{49\sqrt{2}\,\delta}{(r_1+r_2)^2}\cos^2\theta$$

so it suffices to show $h \geq 0$. After isolating the radical, squaring both sides of the resulting equations and simplifying, the required condition can be written

$$\frac{1}{d} + \frac{\delta}{4\cos\theta} - \frac{49\delta^2 \sin^2\theta}{16\bar{r}^2} \geq \left(\bar{r} + \frac{7\delta\cos^2\theta}{2\bar{r}}\right)^2 \tag{16}$$

where $\bar{r} = (r_1 + r_2)/2$.

Now $\bar{r} + \frac{7\delta}{2\bar{r}}$ simplifies to $\sqrt{1+\delta/4}$. Using this, the right-hand side of (16) can be written

$$\left(\sqrt{1+\frac{\delta}{4}} - \frac{7\delta}{2\bar{r}}\sin^2\theta\right)^2.$$

Expanding this out and then moving terms to the right or left of the inequality according to their sign reduces the sufficient condition to

$$\frac{\delta(1-\cos\theta)}{4\cos\theta} + \frac{7\delta\sin^2\theta}{\bar{r}}\sqrt{1+\frac{\delta}{4}} \geq (1-\frac{1}{d}) + \frac{49\delta^2\sin^4\theta}{4\bar{r}^2} + \frac{49\delta^2\sin^2\theta}{16\bar{r}^2}.$$

Writing $1 - 1/d$ as $2\delta\sin^2 t/(d(1+d))$ and then dividing both sides by $\delta\sin^2\theta$ and multiplying by $\bar{r}$ reduces the problem to showing that

$$\frac{\bar{r}}{4\cos\theta(1+\cos\theta)} + 7\sqrt{1+\frac{\delta}{4}} \geq \frac{2\bar{r}}{d(1+d)} + \frac{49\delta\sin^2\theta}{4\bar{r}} + \frac{49\delta}{16\bar{r}}.$$

At this point, the crude estimates $1/2 \leq \bar{r} \leq 1$, $d \geq 1$ and $\delta \leq 4/55$ show that the right hand side is at most

$$1 + \frac{5\cdot 49\delta}{8} = \frac{71}{22} < 4$$

while the left side clearly exceeds 7. This (finally) completes the proof. QED

Theorem 3 shows that the Euler homothetic geodesics are minimal when viewed within the isosceles subsystem of the full three-body problem for masses such that no spiraling occurs near the Euler restpoints. It seems likely that with this assumption, they are still minimal in the planar problem and the three-body problem in $\mathbb{R}^d$, but this is still an open problem. Some evidence for this is provided by local variational studies. It is known that for variations of the action integral, the non-spiraling Euler homothetic orbits have Morse index zero, and so are at least locally minimal [1]. See also [12] where the connection to the spiraling at the Euler restpoints is made clear. Of course it is still a big step to go from local to global.

Acknowledgements

This research was supported by NSF grant DMS-1712656. It was also supported by MSRI, Berkeley, during the semester on Hamiltonian Systems in the Fall of 2019. It benefitted from helpful discussions with R.Montgomery, G.Yu, N.Duignan and C.Jackman.

References

1. V. Barutello and S. Secchi, *Morse index properties of colliding solutions to the N-body problem* Ann. I. H. Poincaré -AN 25 (2008) 539—565.
2. D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol.33, American Mathematical Society, Providence, 2001.
3. A. Chenciner, *Action minimizing solutions of the Newtonian n-body problem : from homology to symmetry*, Proceedings of the International Congress of Mathematics, Vol. III (Beijing, 2002), 279-294, Higner Ed. Press, 2002

4. A. Da Luz and E. Maderna, *On free time minimizers for the newtonian N-body problem. Math. Proc. Camb. Phil. Soc.* **156**, 209–227 (2014).
5. R. Devaney, Triple collision in the planar isosceles three-body problem, Inv.Math. 60, (1980), 249–267.
6. E. Maderna and A. Venturelli, *Globally minimizing parabolic motions in the Newtonian N-body Problem*, Arch. Ration. Mech., 194:283-313, (2009); arxiv: 1502.06278.
7. C. Marchal, *How the minimization of action avoids singularities*, Celestial Mech. Dynam. Astronom. **83**, 325-354 (2002).
8. R. McGehee, *Triple collision in the collinear three-body problem*, Inv. Math, **27** , (1974) 191–227.
9. R. McGehee, *Singularities in classical celestial mechanics*, Proc.Int.Cong.Math., Helsinki, (1978), 827–834.
10. R. Moeckel, *Orbits of the three-body problem which pass infinitely close to triple collision*, Amer. Jour. Math, v. 103, no. 6, (1981), 1323-1341.
11. R. Moeckel, R. Montgomery, and A. Venturelli, *From Brake to Syzygy*, Archive for Rational Mechanics and Analysis; 204(3):1009-1060, (2012).
12. R. Moeckel, R. Montgomery, and H. Sánchez Morgado, *Free Time Minimizers for the Three-Body Problem*, Cel. Mech. Dyn. Astr. 130,28 (2018).
13. B. Percino and H. Sánchez Morgado, *Busemann functions for the N-body problem.* Arch. Rat. Mech. v. 213, no. 3 (2014) 981- 991.
14. C. Simo, *Analysis of triple collision in the isosceles problem*, in "Classical Mechanics and Dynamical Systems," Marcel Dekker, New York, 1980.

School of Mathematics, University of Minnesota, Minneapolis MN 55455
Email address: rick@math.umn.edu