From Brake to Syzygy

Richard Moeckel, Richard Montgomery, Andrea Venturelli

FROM BRAKE TO SYZYGY

RICHARD MOECKEL, RICHARD MONTGOMERY, AND ANDREA VENTURELLI

Abstract. In the planar three-body problem, we study solutions with zero initial velocity (brake orbits). Following such a solution until the three masses become collinear (syzygy), we obtain a continuous, flow-induced Poincaré map. We study the image of the map in the set of collinear configurations and define a continuous extension to the Lagrange triple collision orbit. In addition we provide a variational characterization of some of the resulting brake-to-syzygy orbits and find simple examples of periodic brake orbits.

1. Introduction and Main Results.

This paper concerns the interplay between brake orbits and syzygies in the Newtonian three body problem. A *brake orbit* is a solution, not necessarily periodic, for which the velocities of all three bodies are zero at some instant, the 'brake instant'. Brake orbits have zero angular momentum and negative energy. A *syzygy* occurs when the three bodies become collinear. We will count binary collisions as syzygies, but exclude triple collision.

Lagrange [9] discovered a brake orbit which ends in triple collision. The three bodies form an equilateral triangle at each instant, the triangle shrinking homothetically to triple collision. Extended over its maximum interval of existence, Lagrange's solution explodes out of triple collision, reaches maximum size at the brake instant, and then shrinks back to triple collision.

Lagrange's solution is the only negative energy, zero angular momentum solution without syzygies [16, 17]. In particular, brake orbits have negative energy, and zero angular momentum, and so all of them, *except* Lagrange's, suffer syzygies. Thus we have a map taking a brake initial condition to its first syzygy. We call this map the *syzygy map*. Upon fixing the energy and reducing by symmetries, the domain and range of the map are topologically punctured open discs, the punctures corresponding to the Lagrange orbit. See figure 4 where the map takes the top "zero-velocity surface", or upper Hill boundary, of the solid Hill's region to the plane inside representing the collinear configurations. The exceptional Lagrange orbit runs from the central point on the top surface to the origin in the plane (which corresponds to triple collision), connecting the puncture in the domain to the puncture in the range.

Theorem 1. *The syzygy map is continuous. Its image contains a neighborhood of the binary collision locus. For a large open set of mass parameters, including equal masses, the map extends continuously to the puncture, taking the equilateral triangle of Lagrange to triple collision.*

Date: June 1, 2011 (Preliminary Version).
2000 *Mathematics Subject Classification.* 70F10, 70F15, 37N05, 70G40, 70G60, 70H12.
Key words and phrases. Celestial mechanics, three-body problem, brake orbits, syzygy.

Remark on the Range. Numerical evidence suggests that the syzygy map is not onto (see figure 5). The closure of its range lies strictly inside the collinear Hill's region. A heuristic explanation for this is as follows. The boundary of the domain of the syzygy map is the collinear zero velocity curve, i.e., the collinear Hill boundary. Orbits starting on this curve remain collinear for all time and so are in a permanent state of syzygy. Nearby, non-collinear orbits oscillate around the collinear invariant manifold and take a certain time to reach syzygy, which need not approach zero as the initial point approaches the boundary of the domain. Thus the nearby orbits have time to move away from the boundary before reaching syzygy. It may be that the syzygy map extends continuously to the boundary but we do not pursue this question here.

Collision-free syzygies come in three types, 1, 2, and 3 depending on which mass lies between the other two at syzygy. Listing the syzygy types in temporal order yields the syzygy sequence of a solution. (The syzygy sequence of a periodic collision-free solution encodes its free homotopy type, or braid type.) In ([18], [29]) the notion of syzygy sequence was used as a topological sorting tool for the three-body problem. (See also [19] and [14].) A "stuttering orbit" is a solution whose syzygy sequence has a stutter, meaning that the same symbol occurs twice in a row, as in "11" "22" or '33". For topological and variational reasons, one of us had believed that stuttering sequences were rare. The theorem easily proves the contrary to be true.

Corollary 1. *Within the negative energy, zero angular momentum phase space for the three body problem there is an open and unbounded set corresponding to stuttering orbits.*

Proof of the corollary. If a collinear configuration q is in the image of the syzygy map, and if v is the velocity of the brake orbit segment at q, then by running this orbit backwards, which is to say, considering the solution with initial condition $(q, -v)$, we obtain a brake orbit whose next syzygy is q, with velocity $+v$. This brake orbit is a stuttering orbit as long as q is not a collision point. Perturbing initial conditions slightly cannot destroy stutters, due to transversality of the orbit with the syzygy plane. QED

Periodic Brake Orbits. In 1893 a mathematician named Meissel conjectured that if masses in the ratio 3, 4, 5 are place at the vertices of a 3-4-5 triangle and let go from rest then the corresponding brake orbit is periodic. Burrau [3] reported the conversation with Meissel and performed a pen-and-paper numerical tour-de-force which suggested the conjecture may be false. This "Pythagorean three-body problem" became a test case for numerical integration methods. Szehebely [27] carried the integration further and found the motion ends (and begins) in an elliptic-hyperbolic escape Peters and Szehebeley [28] perturbed away from the Pythagorean initial conditions and with the help of Newton iteration found a periodic brake orbit.

Modern investigations into periodic brake orbits in general Hamiltonian systems began with Seifert's [22] 1948 topological existence proof for the existence of such orbits for harmonic-oscillator type potentials. (Otto Raul Ruiz coined the term "brake orbits" in [21].) We will establish existence of periodic brake orbits in the three-body problem by looking for brake orbits which hit the syzygy plane C orthogonally. Reflecting such an orbit yields a periodic brake orbit. Assume the

masses are $m_1 = m_2 = 1, m_3 > 0$. Then there is an invariant isosceles subsystem of the three-body problem and we will prove:

Theorem 2. *For m_3 in an open set of mass parameters, including $m_3 = 1$, there is a periodic isosceles brake orbit which hits the syzygy plane C orthogonally upon its 2nd hit (see figure 6).*

Do their exist brake orbits, besides Lagrange's whose 1st intersection with C is orthogonal? We conjecture not. Let $I(t)$ be the total moment of inertia of the three bodies at time t. The metric on shape space is such that away from triple collision, a curve orthogonal to C must have $\dot{I} = 0$ at intersection. This non-existence conjecture would then follow from the validity of

Conjecture 1. *$\dot{I}(t) < 0$ holds along any brake orbit segment, from the brake time up to and including the time of 1st syzygy.*

Our evidence for conjecture 1 is primarily numerical. If this conjecture is true then we can eliminate the restriction on the masses in theorem 1. See the remark following proposition 10.

Variational Methods. We are interested in the interplay between variational methods, brake orbits, and syzygies. If the energy is fixed to be $-h$ then the natural (and oldest) variational principle to use is that often called the Jacobi-Maupertuis action principle, described below in section 4. (See also [2], p. 37 eq. (2).) The associated action functional will be denoted A_{JM} (see eq. (30)). Non-collision critical points γ for A_{JM} which lie in the Hill region are solutions to Newton's equations with energy $-h$. Curves inside the Hill region which minimize A_{JM} among all compenting curves in the Hill region connecting two fixed points, or two fixed subsets will be called *JM minimizers*.

We gain understanding of the syzygy map by considering JM-minimizers connecting a fixed syzygy configuration q to the Hill boundary.

Theorem 3. *(i).JM minimizers exist from any chosen point q_0 in the interior of the Hill boundary to the Hill boundary. These minimizers are solutions. When not collinear, a minimizer has at most one syzygy: q_0.*

(ii).There exists a neighborhood $\mathcal{U}$ of the binary collision locus, such that if $q_0 \in \mathcal{U}$, then the minimizers are not collinear.

(iii).If q_0 is triple collision then the minimizer is unique up to reflection and is one half of the Lagrange homothetic brake solution.

Proof of the part of Theorem 1 regarding the image. Let $\mathcal{U}$ be the neighborhood of collision locus given by Theorem 3. If $q_0 \in \mathcal{U}$, the minimizers of Theorem 3 realize non-collinear brake orbits whose first syzygy is q_0, therefore the image of the syzygy map contains $\mathcal{U}$. QED

An important step in the proof of theorem 3 is of independent interest.

Lemma 1. *[Jacobi-Maupertuis Marchal's lemma] Given two points q_0 and q_1 in the Hill region, a JM minimizer exists for the fixed endpoint problem of minimizing $A_{JM}(\gamma)$ among all paths γ lying in the Hill region and connecting q_0 to q_1. Any such minimizer is collision-free except possibly at its endpoint. If a minimizer does not touch the Hill boundary (except possibly at one endpoint) then after reparametrization it is a solution with energy $-h$.*

We can be more precise about minimizers to binary collision when two or all masses are equal. Let r_{ij} denote the distance beween mass i and mass j.

Theorem 4 (Case of equal masses.)**.** *(a) If $m_1 = m_2$ and if the starting point q_0 is a collision point with $r_{12} = 0$ then the minimizers of Theorem 3 are isosceles brake orbits: $r_{13} = r_{23}$ throughout the orbit.*

(b) If $m_1 = m_2$ and if the starting collinear point q_0 is such that $r_{13} < r_{23}$ (resp. $r_{13} > r_{23}$), then a minimizer γ of Theorem 3 satisfies this same inequality: at every point $\gamma(t)$ we have $r_{13}(t) < r_{23}(t)$ (resp. $r_{13}(t) > r_{23}(t)$).

(c) If all three masses are equal, and if q_0 is a collinear point, a minimizer γ of Theorem 3 satisfies the same side length inequalities as q_0 : if $r_{12} < r_{13} < r_{23}$ for q_0, then at every point $\gamma(t)$ of γ we have $r_{12}(t) < r_{13}(t) < r_{23}(t)$.

Part (c) of this theorem suggest:

Conjecture 2. *If three equal masses are let go at rest, in the shape of a scalene triangle with side lengths $r_{12} < r_{13} < r_{23}$ and attract each other according to Newton's law then these side length inequalities $r_{12}(t) < r_{13}(t) < r_{23}(t)$ persist up to the instant t of first syzygy,*

Commentary. Our original goal in using variational methods was to construct the inverse of the syzygy map using JM minimizers. This approach was thwarted due to our inability to exclude or deal with *caustics*: brake orbits which cross each other in configuration space before syzygy. Points on the boundary of the image of the syzygy map appear to be conjugate points – points where non-collinear brake orbits "focus" onto a point of a collinear brake orbit.

Outline and notation. In the next section we derive the equations of motion in terms suitable for our purposes. In section 3.1 we use these equations to rederive the theorem of [16], [17] regarding infinitely many syzygies. We also set up the syzygy map. In section 3.3 we prove theorem 1 regarding continuity of the syzygy map. In section 4 we investigate variational properties of the Jacobi-Maupertuis metric and prove theorems 3 and 4 and the lemmas around them. In section 5 we establish Theorem 2 concerning a periodic isosceles brake orbit.

2. Equation of Motion and Reduction

Consider the planar three-body problem with masses $m_i > 0, i = 1, 2, 3$. Let the positions be $q_i \in I\!R^2 \cong \mathbf{C}$ and the velocities be $v_i = \dot{q}_i \in I\!R^2$. Newton's laws of motion are the Euler-Lagrange equation of the Lagrangian

$$L = K + U \tag{1}$$

where

$$\begin{aligned} 2K &= m_1|v_1|^2 + m_2|v_2|^2 + m_3|v_3|^2 \\ U &= \frac{m_1 m_2}{r_{12}} + \frac{m_1 m_3}{r_{13}} + \frac{m_2 m_3}{r_{23}}. \end{aligned} \tag{2}$$

Here $r_{ij} = |q_i - q_j|$ denotes the distance between the i-th and j-th masses. The total energy of the system is constant:

$$K - U = -h \qquad h > 0.$$

Assume without loss of generality that total momentum is zero and that the center of mass is at the origin, i.e.,

$$m_1 v_1 + m_2 v_2 + m_3 v_3 = m_1 q_1 + m_2 q_2 + m_2 q_3 = 0.$$

Introduce Jacobi variables

$$\xi_1 = q_2 - q_1 \qquad \xi_2 = q_3 - \frac{m_1 q_1 + m_2 q_2}{m_1 + m_2} \tag{3}$$

and their velocities $\dot{\xi}_i$. Then the equations of motion are given by a Lagrangian of the same form (1) where now

$$\begin{aligned} K &= \mu_1 |\dot{\xi}_1|^2 + \mu_2 |\dot{\xi}_2|^2 \\ U &= \frac{m_1 m_2}{r_{12}} + \frac{m_1 m_3}{r_{13}} + \frac{m_2 m_3}{r_{23}}. \end{aligned} \tag{4}$$

The mass parameters are:

$$\mu_1 = \frac{m_1 m_2}{m_1 + m_2} \qquad \mu_2 = \frac{(m_1 + m_2) m_3}{m_1 + m_2 + m_3} = \frac{(m_1 + m_2) m_3}{m} \tag{5}$$

where

$$m = m_1 + m_2 + m_3$$

is the total mass. The mutual distances are given by

$$\begin{aligned} r_{12} &= |\xi_1| \\ r_{13} &= |\xi_2 + \nu_2 \xi_1| \\ r_{23} &= |\xi_2 - \nu_1 \xi_1| \end{aligned} \tag{6}$$

where

$$\nu_1 = \frac{m_1}{m_1 + m_2} \qquad \nu_2 = \frac{m_2}{m_1 + m_2}.$$

2.1. **Reduction.** Jacobi coordinates (3) eliminate the translational symmetry, reducing the number of degrees of freedom from 6 to 4. The next step is the elimination of the rotational symmetry to reduce from 4 to 3 degrees of freedom. This reduction is accomplished by fixing the angular momentum and working in the quotient space by rotations. When the angular momentum is zero, there is a particulary elegant way to accomplish this reduction.

Regard the Jacobi variables ξ as complex numbers: $\xi = (\xi_1, \xi_2) \in \mathbf{C}^2$. Introduce a Hermitian metric on $\mathbf{C}^2$:

$$\langle\langle u, v \rangle\rangle = \mu_1 v_1 \overline{w_1} + \mu_2 v_2 \overline{w_2} \tag{7}$$

If $||v||^2 = \langle\langle v, v \rangle\rangle$ denotes the corresponding norm then the kinetic energy is given by

$$||\dot{\xi}||^2 = \mu_1 |\dot{\xi}_1|^2 + \mu_2 |\dot{\xi}_2|^2 = K$$

while

$$||\xi||^2 = \mu_1 |\xi_1|^2 + \mu_2 |\xi_2|^2 = I \tag{8}$$

is the moment of inertia. We will also use the alternative formula of Lagrange:

$$||\xi||^2 = \frac{1}{m} \left(m_1 m_2 r_{12}^2 + m_1 m_3 r_{13}^2 + m_2 m_3 r_{23}^2 \right) \tag{9}$$

where the distances r_{ij} are given by(6). The real part of this Hermitian metric is a Riemannian metric on $\mathbf{C}^2$. The imaginary part of the Hermitian metric is a nondegenerate two-form on $\mathbf{C}^2$ with respect to which the angular momentum constant ω of the three-body problem takes the form

$$\omega = \mathrm{im} \langle\langle \xi, \dot{\xi} \rangle\rangle.$$

The rotation group $S^1 = SO(2)$ acts on $\mathbf{C}^2$ according to $(\xi_1, \xi_2) \to e^{i\theta}(\xi_1, \xi_2)$. To eliminate this symmetry introduce a new variable

$$r = ||\xi|| = \sqrt{I}$$

to measure the overall size of the configuration and let $[\xi] = [\xi_1, \xi_2] \in \mathbf{CP}^1$ be the point in projective space with homogeneous coordinates ξ. Explicity, $[\xi]$ is an equivalence class of pairs of point of $\mathbf{C}^2 \setminus 0$ where $\xi \equiv \xi'$ if and only if $\xi' = k\xi$ for some nonzero complex constant k. Thus $[\xi]$ describes the shape of the configuration up to rotation and scaling. The variables $(r, [\xi])$ together coordinatize the quotient space $\mathbf{C}^2/S^1$.

Recall that the one-dimensional complex projective space $\mathbf{CP}^1$ is essentially the usual Riemann sphere $\mathbf{C} \cup \infty$. The formula $\alpha([\xi_1, \xi_2]) = \xi_2/\xi_1$ gives a map $\alpha : \mathbf{CP}^1 \to \mathbf{C} \cup \infty$, the standard "affine chart". Alternatively, one has the diffeomorphism $St : \mathbf{CP}^1 \to S^2$ (defined in (33)) to the standard unit sphere $S^2 \subset I\!R^3$ by composing α with the inverse of a stereographic projection map $\sigma : S^2 \to \mathbf{C} \cup \infty$. The space $S = \mathbf{CP}^1$ in any of these three forms will be called the *shape sphere*.

In the papers [16] and [5] the sphere version of shape space was used, and the variables $r, [\xi]$ were combined at times to give an isomorphism $\mathbf{C}^2/S^1 \to I\!R^3$ sending $(r, [\xi]) \mapsto rSt([\xi])$. The projective version of the shape sphere, although less familiar, makes some of the computations below much simpler. Triple collision corresponds to $0 \in \mathbf{C}^2$ and the quotient map $\mathbf{C}^2 \setminus 0 \to (\mathbf{C}^2 \setminus 0)/S^1$ is realized by the map

$$\pi : \mathbf{C}^2 \setminus 0 \to Q = (0, \infty) \times \mathbf{CP}^1; \pi(\xi) = (||\xi||, [\xi]). \tag{10}$$

To write down the quotient dynamics we need a description of the kinetic energy in quotient variables, and so we need a way of describing tangent vectors to $\mathbf{CP}^1$. Define the equivalence relation $\equiv$ by $(\xi, u) \equiv (\xi', u')$ if and only if there are complex numbers k, l with $k \neq 0$ such that $(\xi', u') = (k\xi, ku + l\xi)$. It is easy to see that two pairs are equivalent if and only if $T\pi(\xi, u) = T\pi(\xi', u')$ where $T\pi : T(\mathbf{C}^2 \setminus 0) \to T\mathbf{CP}^1$ is the derivative of the quotient map. Thus an equivalence class $[\xi, u]$ of such pairs represents an element of $T_{[\xi]}\mathbf{CP}^1$, i.e., a *shape velocity* at the shape $[\xi]$. One verifies that the expression

$$||[\xi, \dot{\xi}]||^2 = \frac{\mu_1\mu_2}{||\xi||^4}|\xi_1\dot{\xi}_2 - \xi_2\dot{\xi}_1|^2. \tag{11}$$

defines a quadratic form on tangent vectors at $[\xi]$ and as such is a metric. It is the Fubini-Study metric, which corresponds under the diffeomorphism St to the standard 'round' metric on the sphere, scaled so that the radius of the sphere is $1/2$. We emphasize that in this expression, and in the subsequent ones involving the variables $r, [\xi]$, the variable ξ is to be viewed as a homogeneous coordinate on $\mathbf{CP}^1$ so that the $||\xi|| = (\mu_1|\xi_1|^2 + \mu_2|\xi_2|^2)^{1/2}$ occurring in the denominator is not linked to r, which is taken as an independent variable. Indeed (11) is invariant under rotation and scaling of $\xi, \dot{\xi}$.

We have the following nice formula for the kinetic energy:

Proposition 1. $K = \frac{1}{2}||\dot{\xi}||^2 = \frac{1}{2}\left(\dot{r}^2 + \frac{\omega^2}{r^2} + r^2|||[\xi, \dot{\xi}]|||^2\right).$

We leave the proof up to the reader, or refer to [5] for an equivalent version. Taking $\omega = 0$ gives the simple formula

$$K_0 = \tfrac{1}{2}\dot{r}^2 + \tfrac{1}{2}r^2\frac{\mu_1\mu_2}{||\xi||^4}|\xi_1\dot{\xi}_2 - \xi_2\dot{\xi}_1|^2 \tag{12}$$

By homogeneity, the negative potential energy U can also be expressed in terms of $r, [\xi]$. Set

$$V([\xi]) = ||\xi||\, U(\xi). \tag{13}$$

thus defining V. Equivalently, $V([\xi]) = U(\xi/\|\xi\|)$. Since the right-hand side is homogeneous of degree 0 with respect to ξ, and since U is invariant under rotations, the value of $V([\xi])$ is independent of the choice of representative for $[\xi]$. Clearly we have $U(\xi) = \frac{1}{r}V([\xi])$ for $\xi \in \mathbf{C}^2 \setminus 0$ and $(r, [\xi]) = \pi(\xi)$. The function $V : S \to I\!R$ will be called the *shape potential*.

The function $L_{red} : TQ \to I\!R$ given by

$$L_{red}(r, \dot{r}, [\xi, \dot{\xi}]) = K_0 + \frac{1}{r}V([\xi]) \tag{14}$$

will be called the *reduced Lagrangian*. The theory of Lagrangian reduction [11, 16] then gives

Proposition 2. *Let $\xi(t)$ be a zero angular momentum solution of the three-body problem in Jacobi coordinates. Then $(r(t), [\xi(t)]) = \pi(\xi(t)) \in Q = (0, \infty) \times \mathbf{CP}^1$ is a solution of the Euler-Lagrange equations for the reduced Lagrangian L_{red} on TQ.*

2.2. **The Shape Sphere and the Shape Potential.** Let C be the set of collinear configurations. In terms of Jacobi variables $\xi = (\xi_1, \xi_2) \in C$ if and only if the ratio of ξ_1, ξ_2 is real. The corresponding projective point then satisfies $[\xi] \in \mathbf{RP}^1 \subset \mathbf{CP}^1$. Recall that $\mathbf{RP}^1$ can also be viewed as the extended real line $I\!R \cup \infty$ or as the circle S^1. Thus one can say that the normalized collinear shapes form a circle in the shape sphere, S. Taking the size into account one has $C = I\!R^+ \times S^1 \subset I\!R^+ \times S$. Here and throughout, by abuse of notation, we will write C as the set of collinear states, either viewed before or after reduction by the circle action (so that $C \subset \bar{Q}$), or by reduction by the circle action and scaling (so $C \subset \mathbf{CP}^2$).

The binary collision configurations and the Lagrangian equilateral triangles play an important role in this paper. Viewed in $\mathbf{CP}^1$, these form five distinguished points, points whose homogeneous coordinates are easily found. Setting the mutual distances (6) equal to zero one finds collision shapes:

$$b_{12} = [0, 1] \qquad b_{13} = [1, -\nu_2] \qquad b_{23} = [1, \nu_1]$$

where, as usual, the notation $[\xi_1, \xi_2]$ means that (ξ_1, ξ_2) is a representative of the projective point. Switching to the Riemann sphere model by setting $z = \xi_2/\xi_1$ gives

$$b_{12} = \infty \qquad b_{13} = -\nu_2 \qquad b_{23} = \nu_1.$$

The equilateral triangles are found to be at $[1, l_\pm] \in \mathbf{CP}^1$ or at $l_\pm \in \mathbf{C}$ where

$$l_\pm = \frac{m_1 - m_2}{2(m_1 + m_2)} \pm \frac{\sqrt{3}}{2}\, i = \frac{\nu_1 - \nu_2}{2} \pm \frac{\sqrt{3}}{2}\, i. \tag{15}$$

We will choose coordinates on the shape sphere such that all of these special shapes have simple coordinate representations [16]. In these coordinates, the shape potential will also have a relatively simple form. To carry out this coordinate change, we use the well-known fact from complex analysis that there is a unique

conformal isomorphism (i.e., a fractional linear map) of the Riemann sphere taking any triple of points to any other triple. Thus one can move the binary collisions to any convenient locations. We move them to the third roots unity on the unit circle. Working projectively in homogeneous coordinates, a fractional linear map

$$z = \frac{cw + d}{aw + b}$$

becomes a linear map

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

where $[\xi_1, \xi_2] = [1, z]$ and $[\eta_1, \eta_2] = [1, w]$.

Proposition 3. *Let $\lambda = e^{\frac{2\pi i}{3}}$ and ϕ be the unique conformal map taking $1, \lambda, \overline{\lambda}$ to b_{12}, b_{13}, b_{23} respectively. Then ϕ maps the unit circle to the collinear shapes and 0 and ∞ to the equilateral shapes l_+ and l_-. Moreover, in homogeneous coordinates*

$$\phi([\eta]) = \begin{bmatrix} 1 & -1 \\ l_+ & -l_- \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \tag{16}$$

The proof is routine, aided by the fact that ϕ preserves cross ratios.

We will be using $\eta = (\eta_1, \eta_2)$ as homogeneous coordinates on $\mathbf{CP}^1$ and setting

$$w = \eta_2/\eta_1 \in C.$$

In w coordinates, we have seen that the collinear shapes form the unit circle, with the binary collisions at the third roots of unity and the Lagrange shapes at $w = 0, \infty$. We also need the shape potential in w-variables. It can be calculated from the formulas in the previous subsection by simply setting

$$\xi_1 = \eta_1 - \eta_2 \qquad \xi_2 = l_+\eta_1 - l_-\eta_2 \qquad \eta_1 = 1 \qquad \eta_2 = x + iy. \tag{17}$$

First, (6) gives the remarkably simple expressions

$$\begin{aligned} r_{12}^2 &= (x-1)^2 + y^2 \\ r_{13}^2 &= (x + \tfrac{1}{2})^2 + (y - \tfrac{\sqrt{3}}{2})^2 \\ r_{23}^2 &= (x + \tfrac{1}{2})^2 + (y + \tfrac{\sqrt{3}}{2})^2. \end{aligned} \tag{18}$$

Using these, one can express the norm of the homogeneous coordinates and the shape potential as functions of (x, y). $||\xi||$ is given by (9) and

$$V(x,y) = ||\xi|| \left(\frac{m_1 m_2}{r_{12}} + \frac{m_1 m_3}{r_{13}} + \frac{m_2 m_3}{r_{23}} \right) = \frac{m_1 m_2}{\rho_{12}} + \frac{m_1 m_3}{\rho_{13}} + \frac{m_2 m_3}{\rho_{23}} \tag{19}$$

with $\rho_{ij} = r_{ij}/||\xi||$.

Remark. It is worth saying a bit about the meaning of the expressions eq. (18) and the variables ρ_{ij} occuring in eq (19). A function on $\mathbf{C}^2$ which is homogeneous of degree 0 and rotationally invariant defines a function on $\mathbf{CP}^1$. But the r_{ij} are homogeneous of degree 1, so do not define functions on $\mathbf{CP}^1$ in this simple manner So what is eq. (18) saying? Introduce the local section $\sigma : \mathbf{CP}^1 \setminus \{\infty\} \to \mathbf{C}^2$, given by $[1, w] \to (1, w)$ and the linear map $\tilde{\Phi} : \mathbf{C}^2 \to \mathbf{C}^2$ which induces ϕ. Apply $\tilde{\Phi} \circ \sigma$ to the point $[1, w]$ to form $\tilde{\Phi}(\sigma([1, \eta]) = (\xi_1, \xi_2) \in \mathbf{C}^2$ and then apply the distance functions r_{ij} to this configuration in $\mathbf{C}^2$ to get the r_{ij} of eq. (18). That is, the functions of eq. (18) are $r_{ij} \circ \tilde{\Phi} \circ \sigma$. Then $||\xi||$ in the expression for V is the moment of inertia $I = r^2$ as given by (9) with the r_{ij} there being those given by eq. (18).

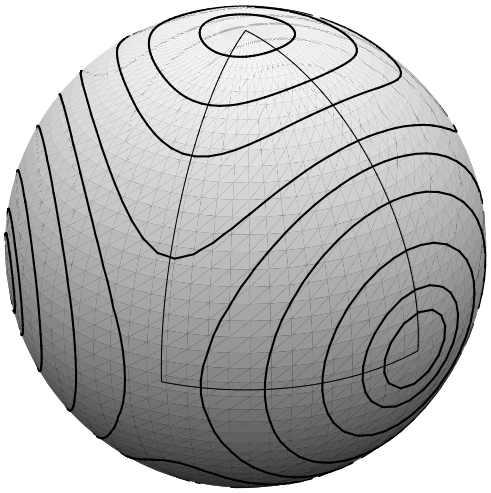

FIGURE 1. Contour plot of the shape potential on the unit sphere $s_1^2+s_2^2+s_3^2 = 1$ in the equal mass case. There is a discrete symmetry of order twelve generated by the reflections in the sides of the indicated spherical triangle.

Alternatively, we can view $\mathbf{CP}^1$ as S^3/S^1 and realize the S^3 by setting $\|\xi\| = 1$. Then the ρ_{ij} are r_{ij} restricted to this S^3, and then understood as S^1 invariant functions.

Figure 1 shows a spherical contour plot of V for equal masses $m_1 = m_2 = m_3$. The equator features the three binary collision singularities as well as three saddle points corresponding to the three collinear or Eulerian central configurations. The equilateral points at the north and south poles of the sphere are the Lagrangian central configurations which are minima of V.

Figure 2 shows contour plots of the shape potential in stereographic coordinates (x, y) for the equal mass case and for $m_1 = 1, m_2 = 2, m_3 = 10$. The unit disk in stereographic coordinates corresponds to the upper hemisphere in the sphere model. When the masses are not equal, the potential is not as symmetric, but due to the choice of coordinates, the binary collisions are still at the roots of unity and the Lagrangian central configuration (which is still the minimum of V) is at the origin.

The variables in figures 1 and 2 are related by stereographic projection:

$$s_1 = \frac{2x}{1+x^2+y^2} \qquad s_2 = \frac{2y}{1+x^2+y^2} \qquad s_3 = \frac{1-x^2-y^2}{1+x^2+y^2} \quad .$$

The following result about the behavior of the shape potential will be useful [16]. Consider the potential in the upper hemisphere (the unit disk in stereographic coordinates). $V(x, y)$ achieves its minimum at the origin. It turns out that V is strictly increasing along radial line segments from the origin to the equator.

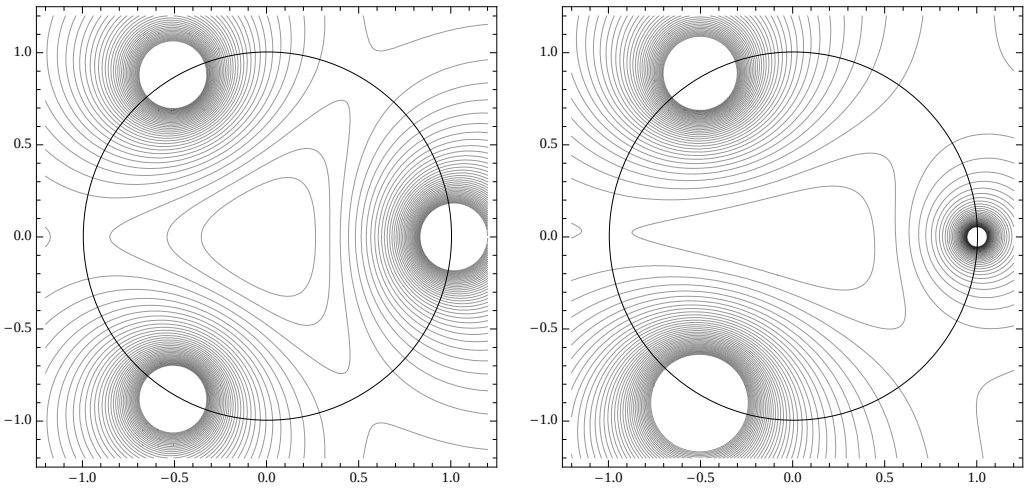

FIGURE 2. Contour plot of the shape potential in stereographic coordinates (x, y). The unit disk corresponds to the upper hemisphere in the sphere model. On the left is the equal mass case as in figure 1. On the right, the masses are $m_1 = 1, m_2 = 2, m_3 = 10$.

Proposition 4. *(Compare with lemma 4, section 6 of* [16]*.) For all positive masses, the shape potential* $V(x, y)$ *satisfies*

$$xV_x + yV_y = \phi(x, y)(1 - x^2 - y^2)$$

where $\phi(x, y) \geq 0$ *with strict inequality if* $(x, y) \neq (0, 0)$.

Proof. Write

$$V = \frac{m_1 m_2}{\rho_{12}} + \frac{m_1 m_3}{\rho_{13}} + \frac{m_2 m_3}{\rho_{23}}$$

with $\rho_{ij} = r_{ij}/||\xi||$, and $r_{ij}, ||\xi||$ expressed as functions of (x, y) using (18) and (9). Then a computation shows that

$$xV_x + yV_y = \phi(x, y)(1 - x^2 - y^2)$$

where

$$\phi = \frac{m_1 m_2 m_3}{2(m_1 + m_2 + m_3)||\xi||}(m_1 g_1 + m_2 g_2 + m_3 g_3) \tag{20}$$

and

$$g_1 = (r_{13}^2 - r_{12}^2)(\frac{1}{r_{12}^3} - \frac{1}{r_{13}^3}) \quad g_2 = (r_{23}^2 - r_{12}^2)(\frac{1}{r_{12}^3} - \frac{1}{r_{23}^3}) \quad g_3 = (r_{23}^2 - r_{13}^2)(\frac{1}{r_{13}^3} - \frac{1}{r_{23}^3}).$$

Note that $g_1 \geq 0$ with strict inequality except on the line where $r_{12} = r_{13}$. Similar properties hold for g_2 and g_3 and the proposition follows. QED

2.3. **Equations of Motion and Hill's Region.** We can derive the equations of motion on $Q = (0, \infty) \times \mathbf{CP}^1$ by calculating the reduced Lagrangian L_{red} in any

convenient coordinates and then writing out the resulting Euler-Lagrange equations. We use the coordinates r, x, y as above with $w = \eta_2/\eta_1 = x + iy$. (See (10, (17) and also eq. (18), (9) and (19).) Then

$$K_0 = \tfrac{1}{2}\dot{r}^2 + \tfrac{1}{2}\kappa(x,y)r^2(\dot{x}^2 + \dot{y}^2) \qquad\qquad \kappa = \frac{3\mu_1\mu_2}{||\xi||^4},$$

where $\|\xi\|^2$ as a function of x, y is obtained by plugging the expressions (18) into Lagrange's identity (9). Then

$$L_{red}(r, x, y, \dot{r}, \dot{x}, \dot{y}) = K_0 + \frac{1}{r}V(x,y)$$

and so the Euler-Lagrange equations are

$$\begin{aligned} \ddot{r} &= -\frac{1}{r^2}V + \kappa r(\dot{x}^2 + \dot{y}^2) \\ (\kappa r^2\dot{x})^{\cdot} &= \frac{1}{r}V_x + \frac{1}{2}\kappa_x r^2(\dot{x}^2 + \dot{y}^2) \\ (\kappa r^2\dot{y})^{\cdot} &= \frac{1}{r}V_y + \frac{1}{2}\kappa_y r^2(\dot{x}^2 + \dot{y}^2). \end{aligned} \tag{21}$$

Conservation of energy gives

$$K_0 - \frac{1}{r}V(x,y) = -h.$$

Remark The expression $\kappa(x,y)(dx^2 + dy^2)$ describes a spherically symmetric metric on the shape sphere. For example, when $m_1 = m_2 = m_3$ one computes that $\kappa = \frac{1}{(1+x^2+y^2)^2}$ which is the standard conformal factor for expressing the metric on the sphere of radius $1/2$ in stereographic coordinates x, y.

These equations describe the zero angular momentum three-body problem reduced to 3 degrees of freedom by elimination of all the symmetries and separated into size and shape variables. An additional improvement is achieved by blowing up the triple collision singularity at $r = 0$ by introducing the time rescaling $' = r^{\frac{3}{2}}\cdot$ and the variable $v = r'/r$ [12]. The result is the following system of differential equations:

$$\begin{aligned} r' &= vr \\ v' &= \tfrac{1}{2}v^2 + \kappa(x'^2 + y'^2) - V \\ (\kappa x')' &= V_x - \tfrac{1}{2}\kappa v x' + \tfrac{1}{2}\kappa_x(x'^2 + y'^2) \\ (\kappa y')' &= V_y - \tfrac{1}{2}\kappa v y' + \tfrac{1}{2}\kappa_y(x'^2 + y'^2). \end{aligned} \tag{22}$$

The energy conservation equation is

$$\tfrac{1}{2}v^2 + \tfrac{1}{2}\kappa(x'^2 + y'^2) - V(x,y) = -rh \tag{23}$$

Note that $\{r = 0\}$ is now an invariant set for the flow, called the *triple collision manifold*. Also, the differential equations for (v, x, y) are independent of r. Call the rescaled time variable s. Since the rescaling is such that $t'(s) = r^{\frac{3}{2}}$, behavior near triple collision that is fast with respect to the usual time may be slow with respect to s. Motion on the collision manifold could be said to occur in zero t-time.

System (22) could be written as a first-order system in the variables (r, x, y, v, x', y'). Fixing an energy $-h < 0$ defines a five-dimensional energy manifold

$$P_h = \{(r, x, y, v, x', y') : r \geq 0,\ (23)\, holds\}.$$

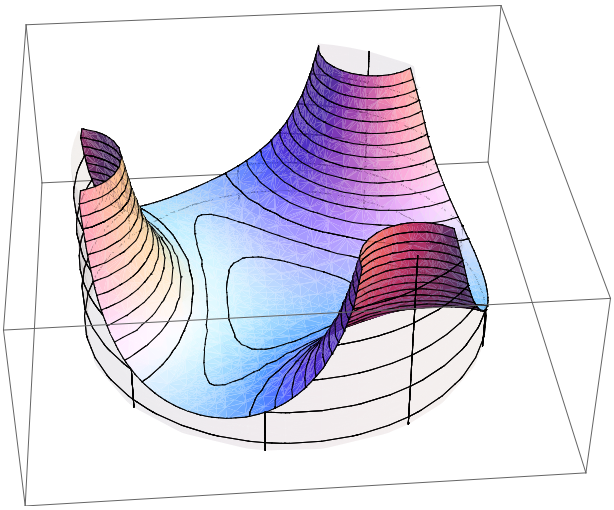

FIGURE 3. Half of the Hill's region for the equal mass three-body problem in stereographic coordinates (x, y, r) where the unit disk corresponds to the upper half of the shape sphere. In these coordinates, the configurations with collinear shapes are in the cylindrical over the unit circle.

The projection of this manifold to configuration space is the *Hill's region*, Q_h. Since the kinetic energy is non-negative the Hill's region is given by

$$Q_h = \{(r, x, y) : 0 \leq r \leq V(x, y)/h\}.$$

Figure 3 shows the part of the Hill's region over the unit disk in the (x, y)-plane for the equal mass case. The resulting solid region has three boundary surfaces. The top boundary surface $r = V(x, y)/h$ is part of the projection to configuration space of the *zero-velocity surface*, the bottom surface, $r = 0$, is contained in the projection of the triple collision manifold. The side walls are part of the vertical cylinder over the unit circle which represents collinear shapes.

2.4. **Visualizing the Syzygy Map.** Figure 4 shows a different visualization of the same Hill's region. This time the shape is viewed as a point $\vec{s} = (s_1, s_2, s_3)$ on the unit sphere which is then scaled by the size variable r to form $r\vec{s}$ and plotted. The region of figure 3 corresponds to the upper half of the solid in the new figure. The collision manifold $r = 0$ is collapsed the origin so that the bottom circle of the syzygy cylinder in figure 3 has been collapsed to a point. The collinear states within the Hill region now form the an unbounded "three-armed" planar region homeomorphic to the interior of the unit disk of figure 3. The syzygy map is the flow-induced map from the top half of the boundary surface in figure 4 to the interior of this three-armed planar region.

Figure 5 illustrates the behavior of the map in the equal mass case, using coordinates which compress the unbounded region C_h into a bounded one. The open upper hemisphere of the shape sphere has been identified with the domain of the syzygy map. The figure shows the numerically computed images of several lines of constant latitude and longitude on the shape sphere. The figure illustrates some of the claims of theorem 1. Note that the image of the map seems to be strictly

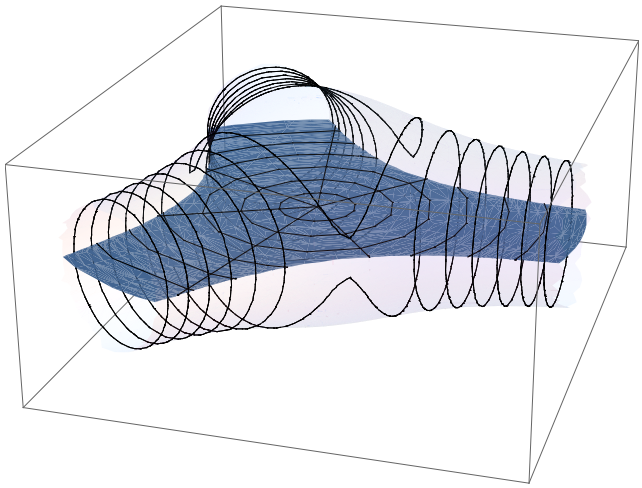

FIGURE 4. The entire Hill's region for the equal mass three-body problem in coordinates $r(s_1, s_2, s_3)$. The syzygy configurations C_h form the planar region, homeomorphic to a disk, dividing the Hill's region in half. The origin represents triple collision (which we count as a syzygy). The syzygy map maps from the upper half of the Hill region's boundary, ∂Q_h^+, to C_h.

smaller than C_h and is apparently bounded by curves connecting the binary collision points on the boundary. It does contain a neighborhood of the binary collision rays however. Also note that a line of high latitude, near the Lagrange homothetic initial condition (the North pole), maps to a small curve encircling the origin. Apparently there is a strong tendency, as yet unexplained, to reach syzygy near binary collision and near the boundary of the image.

2.5. **Flow on the collision manifold: linearization results.** We will need some information about the flow on the triple collision manifold $r = 0$ which we take from [15]. Eq. (23) expresses the triple collision manifold as a two-sphere bundle over the shape sphere, with bundle projection $(0, x, y, v, x', y') \to (x, y)$. The vector field (22) restricted to the triple collision manifold $r = 0$ flow has 10 critical points, coming in pairs $\{p_+, p_-\}$, one pair for each central configuration p. Two of these central configurations correspond to the equilateral triangle configurations of Lagrange and are located in our xy coordinates at the origin and at infinity. $(x, y) = \infty$. We write L for the one at the origin. The other three central configurations are collinear, were found by Euler, and are located on the unit circle in the xy plane, alternating between the three binary collision points. The two equilibria $p_\pm = (x_0, y_0, v_\pm, 0, 0)$ for a given central configuration $p = (x_0, y_0)$ are obtained by solving for v from the $r = 0$ energy equation (23) to get $v = v_\pm = \pm\sqrt{2V(x_0, y_0)}$. The positive square-root v_+ corresponds to solutions 'exploding out" homothetically from that configuration, and the negative square-root v_- with $v = v_- < 0$ corresponds to solutions collapsing into that central configuration. Associated to each central

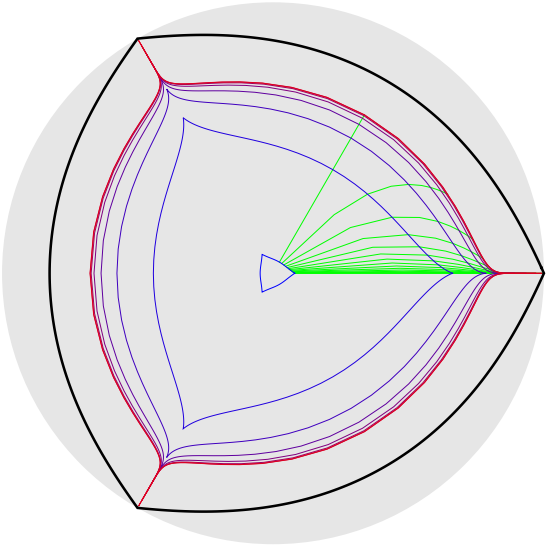

FIGURE 5. Image of the syzygy map in the equal mass case. The open upper hemisphere of the shape sphere has been identified with the domain of the syzygy map. Several circles of constant latitude and a sector of arcs of constant longitude have been followed forward in time to the first syzygy and the resulting image points plotted. The coordinates on the image are $\arctan(r)\,(s_1, s_2)$ so that the horizontal plane in figure 4 is compressed into the shaded open disk. The syzygy configurations C_h now form the region bounded by the heavy black curve. The image of the map, however, seems to be strictly smaller.

configuration we also have the corresponding homothetic solution, which lives in $r > 0$ and forms a heteroclinic connection connecting p_+ to p_-.

We will also need information regarding the stable and unstable manifolds of the equilibria $p_\pm$. This information can be found in [15], Prop 3.3.

Proposition 5. *Each equilibrium $p_\pm = (0, x_0, y_0, v_\pm, 0, 0)$ is linearly hyperbolic. Each has $v_\pm$ as an eigenvalue with corresponding eigenvector tangent to the homothetic solution and so transverse to the collision manifold. Its remaining 4 eigenvectors are tangent to the collision manifold.*

The Lagrange point L_- has a 3 dimensional stable manifold transverse to the collision manifold and a 2-dimensional unstable manifold contained in the collision manifold. The linearized projection of the unstable eigenspace to the tangent space to the shape sphere is onto. At L_+ the dimensions and properties of the stable and unstable manifolds are reversed.

The Euler points E_- each have a 2-dimensional stable manifold transverse to the collision manifold and contained in the collinear invariant submanifold and a 3-dimensional unstable manifold contained in the collision manifold. At E_+, the dimensions and properties of the stable and unstable manifolds are reversed.

Remarks. 1. The $r > 0$ solutions in the stable manifold of L_- are solutions which limit to triple collision in forward time, tending asymptotically to the Lagrange configuration in shape. The final arcs of the solutions of this set can be obtained by minimizing the Jacobi-Maupertuis length between points P and the triple collision point 0 for P varying in some open set of the form $U \setminus C$ where U is a neighborhoodof 0 and C is the collinear set.

2. Because the unstable eigenspace for L_- projects linearly onto the tangent space to the shape sphere it follows that the unstable manifold of L_- cannot be contained in any shape sphere neighborhood $x^2 + y^2 < \epsilon$ of L.

3. The unstable manifold of an exploding Euler equilibrium E_+ lie entirely within the collinear space $z = 0$ and real solutions lying in it form a two-dimensional set of curves. These curves can be obtained by minimizing the Jacobi-Maupertuis length among all *collinear* paths connecting points P and the triple collision point 0 as P varies over collinear configurations in some neighborhood $U \cap C$ of 0.

4. The Sundman inequality implies that $v' \geq 0$ everywhere on the triple collision manifold and that $v(s)$ is strictly increasing except at the 10 equilibrium points. That is, v acts like a Liapanov function on the collision manifold.

3. The Syzygy Map

In this section the syzygy map taking brake initial conditions to their first syzygy will be studied. As mentioned in the introduction, it will be shown that every non-collinear, zero-velocity initial condition in phase space can be followed forward in time to its first syzygy. The goal is to study the continuity and image of the resulting mapping.

3.1. **Existence of Syzygies.** In [16, 17] Montgomery shows that every solution of the zero angular momentum three-body problem, except the Lagrange (equilateral) homothetic triple collision orbit, must have a syzygy in forward or backward time. In forward time, the only solutions which avoid syzygy are those which tend to Lagrangian triple collision. We now rederive this result, using our coordinates.

The result will follow from a study of the differential equation governing the (signed) distance to syzygy in shape space. We take for the signed distance

$$z = 1 - x^2 - y^2$$

where x, y are the coordinates of section 2.3. Note the unit circle $z = 0$ is precisely the set of collinear shapes. From (21), one finds

$$\begin{aligned} \dot{z} &= \frac{2}{\kappa r^2} p_1 \\ \dot{p}_1 &= -\frac{1}{r}(xV_x + yV_y) - r^2(\dot{x}^2 + \dot{y}^2)(\kappa + \tfrac{1}{2}(x\kappa_x + y\kappa_y)). \end{aligned}$$

A computation shows that

$$\kappa + \tfrac{1}{2}(x\kappa_x + y\kappa_y) = \frac{c(1 - x^2 - y^2)}{\|\xi\|^6}$$

where

$$c = \frac{3m_1m_2m_3(m_1m_2 + m_1m_3 + m_2m_3)}{(m_1 + m_2 + m_3)^2}.$$

Using this and proposition 4 gives

$$\begin{aligned} \dot z &= \frac{2}{\kappa r^2} p_1 \\ \dot p_1 &= -F_1(r, x, y, \dot x, \dot y) z \end{aligned} \tag{24}$$

where

$$F_1 = \frac{1}{r}\phi(x,y) + \frac{c r^2(\dot x^2 + \dot y^2)}{\|\xi\|^6}.$$

Note that F_1 is a smooth function for $r > 0$ and satisfies $F_1 \geq 0$ with equality only when $x = y = \dot x = \dot y = 0$.

Using (24) one can construct a proof of Montgomery's result. First note that $z = 0$ defines the syzygy set and $z = p_1 = 0$ is an invariant set, namely the phase space of the collinear three-body problem. Without loss of generality, consider an initial condition with shape (x, y) in the unit disk, i.e., the upper hemisphere in the shape sphere model. Our goal is to show that all such solutions reach $z = 0$.

Remark. The key result of [16] is a differential equation very similar to eq (24) for a variable which was also called z but which we will call z_I now, in order to compare the two. The relation between the current z and this z_I is

$$z_I = \frac{z}{2 - z}$$

and can be derived from the expression $z_I = \frac{1-x^2-y^2}{1+x^2+y^2}$ for the height component of the stereographic projection map $I\!R^2 \to S^2 \setminus (0, 0, -1)$. The important values of these functions on the shape sphere are

$$\text{collinear} : z = 0; z_I = 0$$

$$\text{Lagrange, pos. oriented} : z = 1; z_I = 1$$

$$\text{Lagrange, neg. oriented} : z = +\infty; z_I = -1$$

Proposition 6. *Consider a solution with initial condition lying inside the punctured unit disc in the shape plane : $0 < z(0) < c_1 < 1$, and pointing outward (or at least not inward): $\dot z(0) \leq 0$. Assume that the size of the configuration satisfies $0 < r(t) \leq c_2$ for all time $t \geq 0$, and some positive constant c_2. Then there is a constant $T_0(c_1, c_2) > 0$ and a time $t_0 \in [0, T_0]$ such that $0 < z(t) \leq z(0)$ for $t \in [0, t_0)$ and $z(t_0) = 0$.*

Proof. Consider the projection of the solution to the (z, p_1) plane. By hypothesis, the initial point $(z(0), p_1(0))$ lies in the fourth quadrant of the plane. Since $F_1 \geq 0$, (24) shows that $z(t)$ decreases monotonically on any time interval $[0, t_0]$ such that $z(t) \geq 0$ so $0 \leq z(t) \leq c_1$ holds. An upper bound for t_0 will be now be found.

Let α_1 denote a clockwise angular variable in the (z, p_1) plane. Then (24) gives

$$(z^2 + p_1^2)\dot\alpha_1 = F_1(r, x, y, \dot x, \dot y) z^2 + \frac{2}{\kappa r^2} p_1^2 \geq \frac{1}{r}\phi(x,y) z^2 + \frac{2}{\kappa r^2} p_1^2.$$

On the set where $0 \leq z \leq c_1$ and $0 < r \leq c_2$, the coefficients of z^2 and p_1^2 of the right-hand side each have positive lower bounds. Hence there is a constant $k > 0$ such that $\dot\alpha_1 \geq k$ holds on the interval $[0, t_0]$. It follows that $t_0 \leq \frac{\pi}{2k}$.

It remains to show that the solution actually exists long enough to reach syzygy. It is well-known that the only singularities of the three-body problem are due to collisions. Double collisions are regularizable (and in any case, count as syzygies). Triple collision orbits are known to have shapes approaching either the Lagrangian

or Eulerian central configurations. The Lagrangian case is ruled out by the upper bound on $z(t)$. Eulerian triple collisions can only occur for orbits in the invariant collinear manifold, so this case is also ruled out. QED

The next result gives a uniform bound on time to syzygy for solutions far from triple collision. It is predicated on the well-known fact that if r is large and the energy is negative then configuration space is split up into three disjoint regions, one for each choice of binary pair, and within each region that binary pair moves approximately in a bound Keplerian motion. The approximate period of that motion is obtained from knowledge of the two masses and the percentage of the total energy involved in the binary pair motion. The 'worst' case, i.e. longest period, is achieved by taking the pair to be that with greatest masses, in a parabolic escape to infinity so that all the energy $-h$ is involved in their near Keplerian motion, and thus the kinetic energy of the escaping smallest mass is tending to zero. This limiting 'worst case' period is

$$\tau_* = \frac{1}{(2h)^{3/2}}[\frac{m_i^2 m_j^2}{m_i + m_j}]^{3/2}$$

where the excluded mass m_k is the smallest of the three.

Here is a precise proposition.

Proposition 7. *Let τ_* be the constant above and let β be any positive constant less than 1. Then there is a (small) positive constant $\epsilon_0 = \epsilon_0(\beta)$ such that all solutions with $r(0) \geq \frac{1}{2\epsilon_0}$, energy $-h < 0$, and angular momentum 0 have a syzygy within the time interval $[0, \frac{1}{\beta}\tau_*]$. Moreover, this syzygy occurs before $r = \frac{1}{2}r(0)$, so that $r(t) \geq \frac{1}{\epsilon_0}$ at this syzygy.*

The final sentence of the proposition is added because if initial conditions are such that the far mass approaches the binary pair at a high speed then the perturbation conditions required in the proof will be violated quickly: $r(t)$ will become $O(1)$ in a short time, well before the required syzygy time $\frac{1}{\beta}\tau_*$. In this case the approximate Keplerian frequency of the bound pair is also accordingly high, guaranteeing a syzygy well before perturbation estimates break down and well before the required syzygy time.

Proof. The proof is perturbation theoretic and divides into two parts. In the first part we derive the equations of motions in a coordinate system quite similar to the one which Robinson used to compactify the infinity corresponding to $r \to \infty$ at constant h. In the second part we use these equations to derive the result.

Part 1. Deriving the equations in the new variables. We will use coordinates adapted to studying the dynamics near infinity which are a variation on those introduced by McGehee ([13]) and then modified for the planar three-body problem by Easton, McGehee and Robinson ([6, 7, 20]). Going back to the Jacobi coordinates ξ_1, ξ_2 set

$$\xi_1 = ue^{i\theta}, \xi_2 = \rho e^{i\theta}.$$

thus defining coordinates $(\rho, u, \theta) \in I\!R^+ \times \mathbf{C} \times S^1$. The variables ρ, u coordinatize shape space while θ coordinatizes the overall rotation in inertial space. Then

$$r^2 = \mu_1|u|^2 + \mu_2\rho^2.$$

The reduced kinetic energy (i.e. metric on shape space) is given by

$$K_0 = \frac{\mu_1}{2}|\dot{u}|^2 + \frac{\mu_2}{2}\dot{\rho}^2 - \frac{\mu_1^2}{2r^2}(u \wedge \dot{u})^2.$$

Here

$$u \wedge \dot{u} = u_1\dot{u}_2 - u_2\dot{u}_1 = Im(\bar{u}\dot{u}).$$

(K_0 can be computed two ways: either plug the expressions for the $\dot{\xi}_i$ in terms of $\dot{\rho}, \dot{u}, \dot{\theta}, ...$ into the expression for the kinetic energy and then minimize over $\dot{\theta}$, or plug these same expressions into the reduced metric expression.) Next we make a Levi-Civita transformation by setting $u = z^2$ which gives

$$K_0 = 2\mu_1|z|^2|\dot{z}|^2 + \frac{\mu_2}{2}\dot{\rho}^2 - \frac{2\mu_1^2|z|^4}{r^2}(z \wedge \dot{z})^2.$$

We want to introduce the conjugate momenta and take a Hamiltonian approach. We can write

$$K_0 = 2\mu_1|z|^2\langle A(\rho, z))\dot{z}, \dot{z}\rangle + \frac{1}{2}\mu_2\dot{\rho}^2$$

where A is the symmetric matrix $A = I - B$ with

$$B(\rho, z) = \frac{\mu_1|z|^2}{r^2}\begin{bmatrix} z_2^2 & -z_1z_2 \\ -z_1z_2 & z_1^2 \end{bmatrix}.$$

Here z_1, z_2 denote the real and imaginary parts of z, not new complex variables. The conjugate momenta are

$$\eta = 4\mu_1|z|^2A\dot{z} \qquad y = \mu_2\dot{\rho}.$$

If we write

$$A^{-1} = I + \mu_1 f(\rho, z)$$

we find

$$K_0 = \frac{1}{8\mu_1|z|^2}|\eta|^2 + \frac{1}{2\mu_2}y^2 + \frac{1}{8|z|^2}\langle f(\rho, z)\eta, \eta\rangle.$$

For later use, note that $\mu_1 f = B + B^2 + \dots$ is a positive semi-definite symmetric matrix and $f = O(|z|^4/r^2)$.

The negative of the potential energy is

$$U(\rho, z) = \frac{m_1m_2}{|z|^2} + \frac{m\mu_2}{\rho} + g(\rho, z)$$

where the "coupling term" g is

$$g(\rho, z) \quad = \quad \frac{m_1m_3}{\|\rho + \nu_2z^2\|} + \frac{m_2m_3}{\|\rho - \nu_1z^2\|} - \frac{m\mu_2}{\rho} \quad = O(|z|^4/\rho^3)$$

For $r > r_0$ sufficiently large the Hill region breaks up into three disjoint regions. We are interested for now in the region centered on the 12 binary collision ray. In this case, for $r > r_0$ sufficiently large one finds that $|z|$ is bounded by a constant depending only on the masses and h and the choice of r_0. This bound on $|z|$ tends to a nonzero constant as $r_0 \to \infty$. It then follows from the identity $r^2 = \mu_1\rho^2 + \mu_2|z|^4$ that $r = \sqrt{\mu_2}\rho + O(1)$ for r large.

Next we complete the regularization of the binary collision by means of the the time rescaling $\frac{d}{d\tau} = |z|^2\frac{d}{dt}$. Using the Poincaré trick, the rescaled solutions

with energy $-h$ become the zero-energy solutions of the Hamiltonian system with Hamiltonian function,

$$\begin{aligned}\tilde{H} &= |z^2|(K_0 - U + h)\\ &= \frac{1}{8\mu_1}|\eta|^2 + \frac{|z|^2}{2\mu_2}y^2 + \frac{1}{8}\langle f(\rho,z)\eta,\eta\rangle - m_1m_2 - \frac{m\mu_2|z|^2}{\rho} - g(\rho,z)|z|^2 + h|z|^2.\end{aligned}$$

Computing Hamilton's equations, making the additional substitution

$$x = 1/\rho$$

to move infinity to $x = 0$, and writing $'$ for $d/d\tau$ gives the differential equations:

$$\begin{aligned} x' &= -\frac{|z|^2}{\mu_2}x^2y\\ y' &= -m\mu_2|z|^2x^2 + |z|^2g_\rho - \frac{1}{8}\langle f_\rho\eta,\eta\rangle\\ z' &= \frac{1}{4\mu_1}\eta + \frac{1}{4}f\eta\\ \eta' &= -2(h + \frac{y^2}{2\mu_2} - m\mu_2x - g)z - \frac{1}{8}\langle f_z\eta,\eta\rangle + |z|^2g_z\end{aligned} \tag{25}$$

where the subscripts on f and g denote partial derivatives. These satisfy the bounds:

$$\begin{aligned} f = O(x^2|z|^4) \quad f_z = O(x^2|z|^3) \quad f_\rho = O(x^3|z|^4)\\ g = O(x^3|z|^4) \quad g_z = O(x^3|z|^3) \quad g_\rho = O(x^4|z|^4).\end{aligned}$$

The energy equation is $\tilde{H} = 0$.

Infinity has become $x = 0$, an invariant manifold. At infinity we have $x' = y' = 0$ while other variables satisfy

$$\begin{aligned} z' &= \frac{1}{4\mu_1}\eta\\ \eta' &= -2(h + \frac{y^2}{2\mu_2})z.\end{aligned}$$

Since y is constant this is the equation of a two-dimensional harmonic oscillator.

Part 2. Analysis. Observe that the configuration is in syzygy if and only if the variable u is real . Since $u = z^2$ we have syzygy at time t if and only if $z(t)$ intersects either the real axis or the imaginary axis. If z's dynamics were exactly that of a harmonic oscillator then it would intersect one or the other axis (typically both) twice per period. We argue that z is sufficiently close to an oscillator that these intersections persist.

We begin by establishing uniform bounds on z and η valid for all x sufficiently small. These come from the energy. We rearrange the expression for energy into the form

$$\frac{1}{8\mu_1}|\eta|^2 + \frac{|z|^2}{2\mu_2}y^2 + \frac{1}{8}\langle f(\rho,z)\eta,\eta\rangle + (h - m\mu_2x - g(x,z))|z|^2 = m_1m_2.$$

Choose ϵ_0 so that $x \le \epsilon_0$ implies $|m\mu_2x+g| \le h/2$. By the positive semi-definiteness of f we have

$$\frac{1}{8\mu_1}|\eta|^2 + (h/2)|z|^2 \le m_1m_2$$

which gives our uniform bounds on z, η.

It now follows from equations (25) that for all $\epsilon < \epsilon_0$ and $x \leq \epsilon$ we have

$$\begin{aligned} |x'| &\leq Cx^2|y| \\ |y'| &\leq Cx^2 \\ z' &= \frac{1}{4\mu_1}\eta + O(\epsilon^2) \\ \eta' &= 2H_{12}z + O(\epsilon^2) \end{aligned} \tag{26}$$

where $H_{12} = -h - \frac{y^2}{2\mu_2} + m\mu x + g$, a quantity which represents the energy of the binary formed by masses m_1, m_2. Here C is a constant depending only on the masses, the energy and ϵ_0.

Fix an initial condition x_0, y_0, z_0, η_0 with $x_0 < \epsilon/2$. Write $(x(t), y(t), z(t), \eta(t))$ for the corresponding solution. For C a positive constant, write $\mathcal{R} = \mathcal{R}_{C,\epsilon}$ for the rectangle $0 \leq x \leq \epsilon, |y - y_0| \leq C\epsilon$ in the xy plane. We will show there exists C_1 depending only on the masses, h, on the constant C and y_0 such that the projection $x(t), y(t)$ of our solution lies in $\mathcal{R}_{C,\epsilon}$ for all times t with $|t| \leq 1/C_1\epsilon$. Suppose that the projection of our solution leaves the rectangle $\mathcal{R}$ in some time t. If it first leaves through the y-side, then we have $C\epsilon = |y(t) - y_0| = |\int \dot{y} dt| < C\epsilon^2 \int dt = C\epsilon^2 t$, asserting that $1/\epsilon \leq t$. Thus it takes at least a time $1/\epsilon$ to escape out the y-side. To analyze escape through the x-side at $x = \epsilon$ we enlist Gronwall. Let $C_2(y_0) = C(|y_0| + C\epsilon)$. Compare $x(t)$ to the solution $\tilde{x}$ to $\dot{x} = C_2x^2$ sharing initial condition with $x(t)$, so that $\tilde{x}(0) = x_0$. The exact solution is $\tilde{x} = x_0/(1 - x_0C_2t)$. Gronwall asserts $x(t) \leq \tilde{x}(t)$ as long as $x(t), y(t)$ remain in the rectangle (so that the estimates (26) are valid). But $\tilde{x}(t) \leq \epsilon$ for $t \leq 1/\epsilon C_2$. Consequently it takes our projected solution at least time $t = 1/\epsilon C_2$ to escape out of the x-side, and thus $(x(t), y(t))$ lies within the rectangle for time $|t| \leq 1/\epsilon C_1$, with $1/C_1 = min\{1, 1/C_2(y_0)\}$.

We now analyze the oscillatory part of equations (25). We have

$$H_{12} = -h - \frac{y^2}{2\mu_2} + m\mu x + g = -h - \frac{y^2}{2\mu_2} + O(\epsilon).$$

and let $H^0_{12} = -h - \frac{1}{2}y_0^2$. Then as long as $(x, y) \in \mathcal{R}_{C,\epsilon}$ we have the bound

$$|H^0_{12} - H_{12}(x, y, z, \eta)| = O(\epsilon).$$

holds. Thus the difference between the vector field defining our equations and the "frozen oscillator" approximating equations

$$\begin{aligned} z' &= \frac{1}{4\mu_1}\eta \\ \eta' &= 2H^0_{12}z \end{aligned} \tag{27}$$

which we get by throwing out the $O(\epsilon^2)$ error terms in (26) and replacing H_{12} by H^0_{12} is $O(\epsilon)$.

Now the period of the frozen oscillator is $T_0 = 2\pi\sqrt{2\mu_1}/\sqrt{|H^0_{12}|}$. On the other hand we remain in $\mathcal{R}_{C,\epsilon}$ for at least time $T_{\mathcal{R}} = 1/\epsilon C_1$. Both of these bounds depend on y_0 but we have

$$\frac{T_0}{T_{\mathcal{R}}} = \epsilon \frac{2\pi\sqrt{2\mu_1}\max\{1, C_2(y_0)\}}{\sqrt{h + y_0^2/2}} \leq C_3\epsilon$$

where C_3 does not depend on y_0. Moreover we have a uniform upper bound

$$T_0 \leq T_{\max} = 2\pi\sqrt{2\mu_1}/\sqrt{h}.$$

Hence for $\epsilon_0 > 0$ sufficiently small and $x_0 < \epsilon/2, \epsilon < \epsilon_0$, a solution of (26) remains in $\mathcal{R}_{C,\epsilon}$ for at least time $T_{\max}$ and the difference between the component $z(t)$ of our solution and the corresponding solution $\tilde{z}$ to the linear frozen oscillator equation (27) is of order ϵ in the C^1-norm: C^1 because we also get the $O(\epsilon)$ bound on $\eta = 4\mu_1 z'$. Now, any solution $\tilde{z}$ to the frozen oscillator crosses either the real or imaginary axis, transversally, indeed at an angle of 45 degrees or more, once per half-period. (The worst case scenario is when the oscillator is constrained to a line segment). Thus the same can be said of the real solution $z(t)$ for ϵ_0 small enough: it crosses either the real or imaginary axis at least once per half period.

Finally, the time bounds involving τ_* of the proposition were stated in the Newtonian time. To see these bounds use the inverse Levi-Civita transformation, and note that a half-period of the Levi-Civita harmonic oscillator corresponds to a full period of an approximate Kepler problem with energy H_{12}^0. This approximation becomes better and better as $\epsilon \to 0$. Moreover, the Kepler period decreases monontonically with increasing absolute value of the Kepler energy so that the maximum period corresponds to the infimum of $|H_{12}^0|$ and this is achieved by setting $y = 0$ (parabolic escape), in which case $H_{12}^0 = -h$ and we get the claimed value of τ_* with $\beta = 1$. (Letting $\beta \to 1$ corresponds to letting $\epsilon_0 \to 0$.) QED

These two propositions, Proposition 7 and Proposition 6, lead to Montgomery's result:

Proposition 8. *Consider an orbit with initial conditions satisfying* $0 < z(0) \leq 1$ *and* $r(0) > 0$. *Either the orbit ends in Lagrangian triple collision with* $z(t)$ *increasing monotonically to 1, or else there is some time* $t_0 > 0$ *such that* $z(t_0) = 0$.

Proof. Consider a solution with no syzygies in forward time. By proposition 7 there must be an upper bound on the size: $r(t) \leq c_2$ for some $c_2 > 0$. Next suppose $z(0) = 1$, i.e., the initial shape is equilateral which means $(x(0), y(0)) = (0, 0)$. If $(\dot{x}(0), \dot{y}(0)) = (0, 0)$ the orbit is the Lagrange homothetic orbit which is in accord with the proposition. If $(\dot{x}(0), \dot{y}(0)) \neq (0, 0)$ then for all sufficiently small positive times t_1 one has $0 < z(t_1) < 1$ and $\dot{z}(t_1) < 0$. Then proposition 6 shows that there will be a syzygy. It remains to consider orbits such that $0 < z(0) < 1$.

Suppose $0 < z(0) < 1$ and that the orbit has no syzygies in forward time. It cannot be the Lagrange homothetic orbit and, by the first part of the proof, it can never reach $z(t) = 1$. Proposition 6 then shows that $\dot{z}(t) > 0$ holds as long as the orbit continues to exist, so $z(t)$ is strictly monotonically increasing. By proposition 7 there is a uniform bound $r(t) \leq c_2$ for some c_2. There cannot be a bound of the form $z(t) \leq c_1 < 1$, otherwise a lower bound $\dot{\alpha}_1 \geq k > 0$ as in the proof of proposition 6 would apply. It follows that $z(t) \to 1$.

To show that the orbit tends to triple collision, it is convenient to switch to the blown-up coordinates and rescaled time of equations (22). Since the part of the blown-up energy manifold with $z \geq z(0) > 0$ is compact, the ω-limit set must be a non-empty, compact, invariant subset of $\{z = 1\}$, i.e., $\{(x, y) = (0, 0)\}$. The only invariant subsets are the Lagrange homothetic orbit and the Lagrange restpoints $L_\pm$. Since $W^s(L_+) \subset \{r = 0\}$, the orbit must converge to the restpoint L_- in the collision manfold as $s \to \infty$. In the original timescale, we have a triple collision after a finite time. QED

3.2. **Existence of Syzygies for Orbits in the Collision Manifold.** In the last subsection, it was shown that zero angular momentum orbits with $r > 0$ and not

tending monotonically to Lagrange triple collision have a syzygy in forward time. This is the basic result underlying the existence and continuity of the syzygy map. However, to study the behavior of the map near Lagrange triple collision orbits we need to extend the result to orbits in the collision manifold $\{r = 0\}$. This entails using the equations (22).

Setting $z = 1 - x^2 - y^2$ as before one finds

$$\begin{aligned} z' &= \frac{2}{\kappa} p_2 \\ p_2' &= -(xV_x + yV_y) - (x'^2 + y'^2)(\kappa + \tfrac{1}{2}(x\kappa_x + y\kappa_y)) + \tfrac{1}{2} v p_2. \end{aligned}$$

As in the last subsection, this can be written

$$\begin{aligned} z' &= \frac{2}{\kappa} p_2 \\ p_2' &= -F_2(x, y, x', y') z - \tfrac{1}{2} v p_2 \end{aligned} \tag{28}$$

where

$$F_2 = \phi(x, y) + c(x'^2 + y'^2).$$

F_2 is a smooth function satisfying $F_2 \geq 0$ with equality only when $x = y = x' = y' = 0$. Recall that $'$ denotes differentiation with respect to a rescaled time variable, s.

To see which triple collision orbits reach syzygy, introduce a clockwise angle α_2 in the (z, p_2)-plane. Then

$$(z^2 + p_2^2)\alpha_2' = F_2 z^2 + \frac{2}{\kappa} p_2^2 + \tfrac{1}{2} v z p_2. \tag{29}$$

The cross term in this quadratic form introduces complications and may prevent certain orbits with $r = 0$ from reaching syzygy.

To begin the analysis, consider a case where the cross term is small compared to the other terms, namely orbits with bounded size near binary collision.

Proposition 9. *Let $c_2 > 0$. There is a neighborhood U of the three binary collision shapes and a time $S_0(U, c_2)$ such that if an orbit has shape $(x(s), y(s)) \in U$ and size $r(s) \leq c_2$ for $s \in [0, S_0]$ then there is a (rescaled) time $s_0 \in [0, S_0]$ such that $z(s_0) = 0$.*

Proof. Let δ, τ be the determinant and trace of the symmetric matrix of the quadratic form

$$\begin{bmatrix} F_2 & \frac{1}{4} v \\ \frac{1}{4} v & 2\kappa^{-1} \end{bmatrix}.$$

Iff $\delta, \tau > 0$ then the smallest eigenvalue of this matrix is greater than $2\delta/\tau$, which yields the lower bound

$$\alpha_2' \geq \frac{2\delta}{\tau} > 0.$$

From (18) it follows that near binary collision, one of the three shape variables $r_{ij} \approx 0$ while the other two satisfy $r_{ik}, r_{jk} \approx \sqrt{3}$. Consider, for example, the binary collision $r_{12} = 0$ at $(x, y) = (1, 0)$. One has $F_2 \geq \phi(x, y)$ and $\phi(x, y) \approx C/r_{12}^3$ for some constant $C > 0$. The matrix entry κ^{-1} is bounded and v can be estimates using the energy relation (23)

$$v^2 \leq 2(V(x, y) - rh) = 2m_1 m_2 / r_{12} + O(1)$$

near $(x, y) = (1, 0), r \leq c_2$. Using these estimates in the trace and determinant gives the asymptotic estimate

$$\alpha_2' \geq \frac{2F_2\kappa^{-1} - \frac{1}{16}v^2}{F_2 + 2\kappa^{-1}} \approx 2\kappa^{-1} > 0.$$

Similar analysis near the other binary collision points yields a neighborhood U in which there is a positive lower bound for α_2'. This forces a syzygy ($z = 0$) in a bounded rescaled time, as required. QED

This result, together with some well-known properties of the flow on the triple collision manifold leads to a characterization of possible triple collision orbits with no syzygy in forward time.

Proposition 10. *Consider an orbit with initial condition $r(0) = 0$ and $0 < z(0) \leq 1$. Either the orbit tends asymptotically to one of the restpoints on the collision manifold or there is a (rescaled) time $s_0 > 0$ such that $z(s_0) = 0$.*

Proof. Recall from section 2.5 that the flow on the triple collision manifold is gradient-like with respect to the variable v, i.e., $v(s)$ is strictly increasing except at the restpoints. It follows that every solution which is not in the stable manifold of one of the restpoints satifies $v(s) \to \infty$. In this case, the energy equation (23) shows that the shape potential $V(x(s), y(s)) \to \infty$ so the shape must approach one of the binary collision shapes. For such orbits, proposition 9 gives a syzygy in a bounded rescaled time and the proposition follows. QED

We will also need a result analogous to proposition 6. Consider an orbit in $\{r = 0\}$ with initial conditions satifying $0 < z(0) \leq c_1 < 1$ and $z'(0) \leq 0$. It will be shown that, under certain assumptions on the masses, every such orbit has $z(s_0) = 0$ at some time $s_0 > 0$.

On any time interval $(0, s_0]$ such that $z(s) > 0$ one has $z'(s) < 0$ and hence $z(s)$ is monotonically decreasing. This follows from the "convexity" condition that $z'' < 0$ whenever $0 < z < 1$ and $z' = 0$ which is easily verified from (28). If the orbit does not reach syzygy in forward time, then $0 < z(s) \leq c_1$ for all $s \geq 0$. It is certainly not in the stable manifold of one of the Lagrangian restpoints at $z = 1$ so, by proposition 10, it must be in the stable manifold of one of the Eulerian (collinear) restpoints at $z = 0$.

It will now be shown that, for most choices of the masses, even orbits in these stable manifolds have syzygies. Let $e_j = (x_j, y_j), j = 1, 2, 3$ denote the collinear central configuration with mass m_j between the other two masses on the line. The two corresponding restpoints on the collision manifold are E_{j-} and E_{j+} with coordinates $(r, x, y, v, x', y') = (0, x_j, y_j, \pm\sqrt{2V(x_j, y_j)}, 0, 0)$. The two-dimensional manifolds $W^s(E_{j-})$ and $W^u(E_{j+})$ are contained in the collinear invariant submanifold. In particular, an orbit with $z(0) > 1$ cannot converge to E_{j-} in forward time. On the other hand, $W^s(E_{j+})$ is three-dimensional and its intersection with the collision manifold is two-dimensional. These are the orbits which might not reach syzygy.

Remark. Conjecture 1 asserts that all orbits reach syzygy with $v < 0$. So, if the conjecture is valid then convergence to E_{j+} would be impossible and the discussion to follow, and the restriction on the masses in the theorem, would be unnecessary.

For most choices of the masses, the two stable eigenvalues of E_{j+} with eigenspaces tangent to the collision manifold are non-real. See [15]. In this case we will say that

E_{j+} is *spiraling*. The spiraling case is more common: real eigenvalues occur only when the mass m_j is much larger than the other two masses. For all masses, at least two of the three Eulerian restpoints E_{j+} are spiraling and for a large open set of masses where no one mass dominates, all of the Eulerian restpoints are spiraling.

Proposition 11. *Let E_{j+} be a spiraling Eulerian restpoint. Then there is a neighborhood U of E_{j+}, and a time $S_0(U)$ such that any non-collinear orbit with initial condition in the local stable manifold $W^s_U(E_{j+})$ has a syzygy in every time interval of length at least S_0.*

Proof. Introduce local coordinates in the energy manifold near E_{j+} of the form (r, a, b, z, p_2) where (a, b) are local coodinates in the collinear collision manifold (the intersection of the invariant collinear manifold with $\{r = 0\}$) and (z, p_2) are the variables of (28). The invariance of the collinear manifold $z = p_2 = 0$ implies that the linearized differential equations for (z, p_2) take the form

$$\begin{bmatrix} z \\ p_2 \end{bmatrix}' = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} z \\ p_2 \end{bmatrix}$$

where the eigenvalues of the matrix are the non-real eigenvalues at the restpoint. The spiraling assumption implies that the angle α_2 in the (z, p_2)-plane satisifies $\alpha_2' > k > 0$ for some constant k. So for the full nonlinear equations, one has $\alpha_2' > k/2 > 0$ in some neighborhood U. Since non-collinear orbits in the local stable manifold remain in U and have nonzero projections to the (z, p_2)-plane, the proposition holds with $S_0 = 2\pi/k$. QED

A mass vector (m_1, m_2, m_3) will said to satisfy the *spiraling assumption* if all of the Eulerian restpoints are spiraling. The next result follows from propositions 10 and 11.

Proposition 12. *Consider an orbit in $\{r = 0\}$ with initial conditions satifying $0 < z(0) \le c_1 < 1$ and $z'(0) \le 0$. Suppose that the masses satisfy the spiraling assumption. Then there is a (rescaled) timetime $s_0 > 0$ such that $z(s_0) = 0$. Moreover $z(s)$ is monotonically decreasing on $[0, s_0]$*

3.3. **Continuity of the Syzygy Map.** In this section, we will prove the statement in Theorem 1 about continuity of the syzygy map and its extension to the Lagrange homothetic orbit.

We begin by viewing the syzygy map in blow-up coordinates (r, x, y, v, x', y'). Recall the notations P_h for the energy manifold and Q_h for the Hill's region. Points on the boundary surface ∂Q_h of the Hill's region can be uniquely lifted to zero-velocity (brake) initial conditions in P_h. Let ∂Q_h^+ be the subset of the boundary with shapes in the open unit disk (which corresponds to the open upper hemisphere in the shape sphere model). This is the upper boundary surface in figure 3. It will be the domain of the syzygy map.

To describe the range, let $\tilde{C}_h = \{(r, x, y, v, x', y') : r > 0, x^2 + y^2 = 1,\ (23)\ holds\}$ be the subset of the energy manifold whose shapes are collinear and let $C_h \subset Q_h$ be its projection to the Hill's region, i.e., the set of syzygy configurations having allowable energies. C_h is the cylindrical surface over the unit circle in figure 3 (the unit circle in $\{r = 0\}$ is the blow-up of the puncture at the origin). The first version of the syzygy map will be a map from part of ∂Q_h^+ to C_h.

Recall that the origin $(x, y) = (0, 0)$ represents the equilateral shape. The corresponding point $p_0 \in \partial Q_h^+$ lifts to a brake initial condition in P_h which is on the

Lagrange homothetic orbit. This orbit converges to the Lagrange restpoint without reaching syzygy. It turns out that p_0 is the only point of ∂Q_h^+ which does not reach syzygy and this leads to the first version of the syzygy map.

Proposition 13. *Every point of $\partial Q_h^+ \setminus p_0$ determines a brake orbit which has a syzygy in forward time. The map $F : \partial Q_h^+ \setminus p_0 \to C_h$ determined by following these orbits to their first intersections with $\tilde{C}_h$ and then projecting to C_h is continuous.*

Proof. The surface ∂Q_h intersects the line $x = y = 0$ transversely at p_0. So every point of $\partial Q_h^+ \setminus p_0$ satisfies $0 < z < 1$ where $z = 1 - x^2 - y^2$ as before. Also $r > 0$ on the whole surface ∂Q_h.

Consider the brake initial condition corresponding to such a point. Since all the velocities vanish one has $\dot{z}(0) = 0$. As in the first paragraph of the proof of proposition 6, one finds that $z(t)$ is decreasing as long as $z(t) \geq 0$. In particular, $z(t)$ does not monotonically increase toward 1. It follows from proposition 8 that there is a time $t_0 > 0$ such that $z(t_0) = 0$. In other words, the forward orbit reaches $\tilde{C}_h$.

To see that the flow-defined map $\tilde{F} : \partial Q_h^+ \setminus p_0 \to \tilde{C}_h$ is continuous, note that if the first syzygy is not a binary collision then $\dot{z}(t_0) < 0$. This follows since $z(t)$ is decreasing and since the orbit does not lie in the invariant collinear manifold with $z = \dot{z} = 0$. So the orbit meets $\tilde{C}_h$ transversely. After regularization of double collisions, even orbits whose first syzygy occurs at binary collision can be seen as meeting $\tilde{C}_h$ transversely as in the proof of proposition 7. It follows from transversality that $\tilde{F}$ is continuous. Composing with the projection to the Hill's region shows that F is also continuous. QED

To get a continuous extension of F to p_0, we have to collapse the triple collision manifold back to a point. One way to do this is to replace the coordinates (r, x, y) with $(X, Y, Z) = r(s_1, s_2, s_3)$ where s_i are given by inverse stereographic projection. The Hill's region is shown in figure 4. In this figure, ∂Q_h^+ is the open upper half of the boundary surface and C_h is the planar surface inside (minus the origin). In this model, triple collision has been collapsed to the origin $(X, Y, Z) = (0, 0, 0)$. It is natural to extend the syzygy map by mapping the Lagrange homothetic point p_0 to the triple collision point, i.e., by setting $F(p_0) = (0, 0, 0)$. The extension maps into $\bar{C}_h = C_h \cup 0$.

Theorem 5. *If the masses satisfy the spiraling assumption then the extended syzygy map $F : \partial Q_h^+ \to \bar{C}_h$ is continuous. Moreover, it has topological degree one near p_0.*

Proof. To prove continuity at p_0 it suffices to show that points in ∂Q_h^+ near p_0 have their first syzygies near $r = 0$. The proof will use the blown-up coordinates (r, x, y) and the rescaled time variable s.

Let L_- denote the Lagrange restpoint on the collision manifold to which the orbit of p_0 converges. Let $W^{s+}(L_-), W^{u+}(L_-)$ be the local stable and unstable manifolds of L_- where *local* means that the orbits converge to L_- while remaining in $\{z > 0\}$. It follows from proposition 12 that $z(s)$ is monotonically increasing to 1 along orbits in $W^{s+}(L_-) \cap \{r = 0\}$. A similar argument with time reversed shows that $z(s)$ monotonically decreases from 1 along orbits in $W^{u+}(L_-)$. These last orbits lie entirely in $\{r = 0\}$) by proposition 5.

The unstable manifold $W^{u+}(L_-)$ is two-dimensional and its projection to the (x, y) plane is a local diffeomorphism near L_-. Let D^u be a small disk around L_-

in $W^{u+}(L_-)$. It follows from proposition 12 that all of the points in $D^u \setminus L_-$ can be followed forward to meet $\tilde{C}_h$ transversely. Moreover the monotonic decrease of $z(s)$ implies that if this flow-induced mapping is composed with the projection to C_h and then to the unit circle, the resulting map from the punctured disk to the circle will have degree one.

Let $\gamma^u = \partial D^u$ be the boundary curve of such an unstable disk. Since the unstable manifold is contained in the collision manifold, the first-syzygy map takes γ^u into $\tilde{C}_h \cap \{r = 0\}$. Transversality implies that the first syzygy map is defined and continuous near γ^u. Hence, given any $\epsilon > 0$ there is a neighborhood U of γ_u such that initial conditions in U have their first syzgies with $r < \epsilon$. Standard analysis of the hyperbolic restpoint L_- then shows that any orbit passing sufficiently close to L_- will exit a neighborhood of L_- through the neigborhood U of γ^u and therefore reaches its first syzygy with $r < \epsilon$. We have established the continuity of the syzygy map at p_0.

Now consider a small disk, D, around p_0 in ∂Q_h^+. It has already been shown that every point in $D \setminus p_0$ can be followed forward under the flow to meet $\tilde{C}_h$ transversely with $z(s)$ decreasing monotonically. If D is sufficiently small it will follow the Lagrange homothetic orbit close enough to L_- for the argument of the previous paragraph to apply. In other words the syzygy map takes $D \setminus p_0$ into $\{r < \epsilon\}$ which implies continuity of the extension at p_0. As before, the monotonic decrease of $z(s)$ implies that the map from the punctured disk to the unit circle has degree one. Hence the map with triple collision collapsed to the origin has local degree one. QED

This completes the proof of the continuity statements in Theorem 1 and also shows that the extended map covers a neighborhood of the triple collision point. See figure 5 for a picture illustrating this theorem in the equal mass case.

4. Variational methods: Existence and regularity of minimizers

In this section, we prove theorems 3, 4 and lemmas 1 by applying the direct method of the Calculus of Variations to the Jacobi-Maupertuis [JM] action:

$$A_{JM}(\gamma) = \int_{t_0}^{t_1} \sqrt{K_0}\sqrt{2(U-h)}\, dt. \tag{30}$$

Here $\gamma : [t_0, t_1] \to Q_h$ is a curve. Curves which minimize A_{JM} within some class of curves will be called "JM minimizers".

The integrand of the functional A_{JM} is homogeneous of degree one in velocities and so is independent of how γ is parameterized. The natural domain of definition of the functional is the space of rectifiable curves in Q_h. We recall some notions about Fréchet rectifiable curves. See [8] for more details. Kinetic energy K_0 induces a complete Riemannian metric, denoted $2K_0$ on the manifold with boundary Q_h. We denote by d_0 the associated Riemannian distance. The Fréchet distance between two continuous curves $\gamma : [t_0, t_1] \to Q_h$ and $\gamma' : [t_0', t_1'] \to Q_h$ is defined to be

$$\rho(\gamma, \gamma') = \inf_h \sup_{t \in [t_0, t_1]} d_0(\gamma(t), \gamma'(h(t))),$$

where the infimum is taken over all orientation preserving homeomorphisms $h : [t_0, t_1] \to [t_0', t_1']$. We consider two curves equivalent if the Fréchet distance between them is zero. An equivalence class of such curves can be seen as an unparametrized

curve. The Fréchet distance is a complete metric on the set of unparametrized curves. The length $\mathcal{L}(\gamma)$ of a parametrized curve $\gamma : [t_0, t_1] \to Q_h$ can be defined to be the supremum of the lengths of broken geodesics associated to subdivisions of $[t_0, t_1]$. This length equals the usual $2K_0$ length when the curve is C^1. Equivalent curves have equal lengths. A curve is called rectifiable if its length is finite. For a generic parametrization of a rectifiable curve, the integral (30) must be interpeted as a Weierstrass integral (see [8]). This integral is independent of parameterization.

We will use the following two facts in our proof of Theorem 3. If $K \subset Q_h$ is compact, then the set of curves in Q_h which intersect K and have length bounded by a given constant forms a compact set in the Fréchet topology. (This fact is a theorem attributed to Hilbert.) The second fact asserts that both the length functional $\mathcal{L}$ and the JM action functional A_{JM} are lower semicontinuous in the Fréchet topology. (Again, see [8]).

Since the Jacobi-Maupertuis action degenerates on the Hill boundary we will need to construct a tubular neighborhood of the Hill boundary within which we can characterize JM-minimizers to the Hill boundary. Our construction follows [22].

Proposition 14. *There exists $\epsilon > 0$, a neighborhood S_h^ϵ of ∂Q_h (in Q_h), an analytic diffeomorphism $\Phi : \partial Q_h \times [0, \epsilon] \to S_h^\epsilon$ (satisfying $\Phi(\partial Q_h, 0) = \partial Q_h$) and a strictly positive constant M, such that if $\delta \in (0, \epsilon)$, and $x \in \partial Q_h$, the curve $y \mapsto \Phi(x, y),\ y \in [0, \delta]$ is a reparametrization of an arc of the brake solution starting from x, and its length is smaller or equal to $M\delta$. If $q = \Phi(x, \delta)$ then for every rectifiable curve γ in Q_h joining q to ∂Q_h we have*

$$A_{JM}(\gamma) \geq \delta^{3/2},$$

with equality if and only if γ is obtained by pasting $(\Phi(x, y))_{y \in [0,\delta]}$ to any arc contained in ∂Q_h. Moreover, there exists a strictly positive constant α, independent from δ, such that $U \geq h + \alpha\delta$ on $Q_h \setminus S_h^\delta$, where we term $S_h^\delta = \Phi(\partial Q_h, [0, \delta])$.

We postpone the proof of this Proposition, and instead use it now to prove Theorem 3 and Lemma 1. We will say that S_h^ϵ is a Seifert tubular neighborhood of the Hill boundary, and that $\Phi(\partial Q_h, \epsilon)$ is the inner boundary of S_h^ϵ. We prove now the first part of Theorem 3.

Proposition 15. *Given a point q_0 in the interior of the Hill region Q_h, there exist a JM action minimizer among rectifiable curves joining q_0 to the Hill boundary ∂Q_h.*

Proof. Let S_h^ϵ be the Seifert tubular neighborhood given by Proposition 14. If $q_0 \in S_h^\epsilon$ then the unique brake orbit joining q_0 to the Hill boundary is a JM-minimizer.

If $q_0 \notin S_h^\epsilon$, let γ_n be a JM minimizing sequence of rectifiable curve joining q_0 to ∂Q_h. Without loss of generality we can assume all these curves are Lipschitz and defined on the unit interval $[0, 1]$. Let $t_n \in (0, 1)$ be the first time such that $\gamma_n(t)$ touches the inner boundary $\Phi(\partial Q_h, \epsilon)$ of S_h^ϵ. Let us replace $\gamma\big|_{[t_n,1]}$ by the unique brake solution joining $\gamma_n(t_n)$ to ∂Q_h. By Proposition 14, this modification decreases the JM action of γ_n, so our sequence is still a minimizing one. Let C be an upper bound for the numbers $A_{JM}(\gamma_n)$. Applying Proposition 14 again we get

$$C \geq A_{JM}(\gamma_n) \geq \epsilon^{3/2} + \mathcal{L}(\gamma_n\big|_{[0,t_n]})\sqrt{\alpha\epsilon}, \tag{31}$$

therefore

$$\mathcal{L}(\gamma_n) \leq \frac{C - \epsilon^{3/2}}{\sqrt{\alpha\epsilon}} + M\epsilon, \tag{32}$$

for all n, proving that the lengths $\mathcal{L}(\gamma_n)$ are bounded. Since $\gamma_n(0) = q_0$ for all n, Hilbert's theorem discussed above applies: the sequence of curves γ_n is relatively compact in the Fréchet topology. Therefore a subsequence of this sequence converges to a rectifiable curve γ in Q_h. This γ is a JM minimizer by the lower semicontinuity of the JM action. QED

Let $d_{JM}(q_0, q_1)$ denote the infimum of the JM-action among rectifiable curves in Q_h joining q_0 to q_1, and let $d_{JM}(q_0, \partial Q_h)$ denote the minimum of the JM-action among rectifiable curves in Q_h joining q_0 to the Hill boundary.

Lemma 1, the JM Marchal lemma, proof.

Proof. First we establish the existence of a minimizer.

If $d_{JM}(q_0, q_1) \geq d_{JM}(q_0, \partial Q_h) + d_{JM}(q_1, \partial Q_h)$, then take a curve realizing the minimum in $d_{JM}(q_0, \partial Q_h)$, another curve realizing the minimum in $d_{JM}(q_1, \partial Q_h)$ and join these two curves by any curve lying on the Hill boundary and connecting the endpoints of these two. In this way we get a rectifiable curve γ joining q_0 to q_1 and (possibly) spending some time on the Hill boundary. Since the JM action of any curve on the Hill boundary is zero, we have that $A_{JM}(\gamma) = d_{JM}(q_0, \partial Q_h) + d_{JM}(q_1, \partial Q_h)$. Hence γ is our minimizer. (As a bonus we have shown that $d_{JM}(q_0, q_1) = d_{JM}(q_0, \partial Q_h) + d_{JM}(q_1, \partial Q_h)$ in this case.)

If $d_{JM}(q_0, q_1) < d_{JM}(q_0, \partial Q_h) + d_{JM}(q_1, \partial Q_h)$, let γ_n be a JM-minimizing sequence of rectifiable curves in Q_h joining q_0 to q_1. We show now that there exists $\epsilon > 0$ such that for n is sufficiently great, the γ_n do not intersect the Seifert tubular neighborhood S_h^ϵ. Assume, for the sake of contradiction that there in fact exists a decreasing sequence of strictly positive real numbers ϵ_n such that $\epsilon_n \to 0$ as $n \to +\infty$ and a subsequence of the minimizing sequence, still denoted $(\gamma_n : [0,1] \to Q_h)_n$, and a sequence of times t_n such that for every n, $\gamma_n(t_n)$ is in $S_h^{\epsilon_n}$. Let $\tilde{\gamma}_n$ be the curve defined in the following way. Follow γ_n from q_0 to $\gamma_n(t_n)$. Join $\gamma_n(t_n)$ to ∂Q_h by the unique brake solution in $S_h^{\epsilon_n}$ with one end point $\gamma_n(t_n)$. Return along the same brake solution, to $\gamma_n(t_n)$. Then continue along γ_n from $\gamma_n(t_n)$ to q_1. The curve $\tilde{\gamma}_n$ joins q_0 to q_1 is rectifiable and touches ∂Q_h, hence $A_{JM}(\tilde{\gamma}_n) \geq d_{JM}(q_0, \partial Q_h) + d_{JM}(q_1, \partial Q_h)$. By proposition 14 we have

$$d_{JM}(q_0, \partial Q_h) + d_{JM}(q_1, \partial Q_h) - 2\epsilon_n^{3/2} \leq A_{JM}(\tilde{\gamma}_n) - 2\epsilon_n^{3/2} \leq A_{JM}(\gamma_n).$$

Taking the limit $n \to +\infty$ we get $d_{JM}(q_0, \partial Q_h) + d_{JM}(q_1, \partial Q_h) \leq d_{JM}(q_0, q_1)$, a contradiction. We have established our desired $\epsilon > 0$ with the property that none of the curves γ_n (n sufficiently large) intersect S_h^ϵ.

Let $C > 0$ be an upper bound of $A_{JM}(\gamma_n)$ and $\alpha = \alpha(\epsilon)$ the positive constant given by Proposition 14. A computation similar to (31) and (32) gives

$$\mathcal{L}(\gamma_n) \leq \frac{C}{\sqrt{\alpha\epsilon}}.$$

Hilbert's Theorem again yields a curve γ in Q_h joining q_0 to q_1 which the a subsequence of the γ_n converges to in the Fréchet topology. By lower semicontinuity of A_{JM}, the curve γ is a JM-minimizer.

Finally, we show that any minimizer $\gamma : [t_0, t_1] \to Q_h$ is collision-free and that every subarc of γ, upon being reparametrized, is a true solution to Newton's equation. The classical Lagrangian action on the reduced space Q is defined by

$$\mathcal{A}_{red}(\sigma) = \int_{t_0}^{t_1} L_{red}(\sigma(t), \dot{\sigma}(t))dt$$

where $\sigma : [t_0, t_1] \to Q$ is any absolutely continuous path and the reduced Lagrangian L_{red} is defined in (14). If $t \in [t_0, t_1]$ is such that $\gamma(t) \in \partial Q_h$, obviously $\gamma(t)$ is not a collision, hence it suffices to prove the statement for subarcs $\gamma\,\big|_{[a,b]}$, $[a, b] \subset [t_0, t_1]$ of γ which never touch the Hill boundary. Let $[a, b]$ be such an interval. By the inequality $A^2 + B^2 \geq 2AB$ applied to the factors $A = \sqrt{K_0}, B = \sqrt{U - h}$ of the integrand for A_{JM} we see that

$$\sqrt{2}A_{JM}\left(\gamma\,\big|_{[a,b]}\right) \leq \mathcal{A}_{red}(\gamma\,\big|_{[a,b]}) - h(b - a)$$

with equality if and only if the energy $K_0 - U$ evaluated on γ is equal to $-h$ for almost all $t \in [a, b]$. We may assume, without loss of generality, that γ is parametrized by its $2K_0$-arclength parameter s so that in the integrand $K_0(s) = 1/2$ for almost every $s \in [a, b]$. Since γ is a JM-minimizer, we know that $\gamma(a) \neq \gamma(b)$, that the action $A_{JM}(\gamma\,\big|_{[a,b]})$ is finite, and that the set of $s \in [a, b]$ such that $\gamma(s)$ is a collision is a closed set of zero Lebesgue measure. Let $\sigma : [\alpha, \beta] \to Q_h$ be the reparametrization of $\gamma\,\big|_{[a,b]}$ defined by $\sigma(t) = \gamma(s(t))$, where $s = s(t)$ is the inverse function of

$$t = t(s) = \int_a^s \frac{dv}{\sqrt{2(U(\gamma(v)) - h)}}, \qquad s \in [a, b].$$

The function $t = t(s)$ is $\mathcal{C}^1$, strictly increasing and $\frac{dt}{ds}(s) > 0$ on an open set with full measure, hence the inverse function $s = s(t)$ is absolutely continuous, and since $s \mapsto \gamma(s)$ is Lipschitz, $t \mapsto \sigma(t) = \gamma(s(t))$ is also absolutely continuous. A simple computation shows that the energy function $K_0 - U$ of σ is the constant $-h$ for almost every $t \in [\alpha, \beta]$. Hence

$$\sqrt{2}A_{JM}\left(\gamma\,\big|_{[a,b]}\right) = \sqrt{2}A_{JM}(\sigma) = \mathcal{A}_{red}(\sigma) - h(\beta - \alpha).$$

But $\gamma\,\big|_{[a,b]}$ is a minimizer of A_{JM}! So this equation says that σ minimizes $\mathcal{A}_{red}$ among all absolutely continuous paths in Q_h joining $\sigma(\alpha)$ to $\sigma(\beta)$ in time $\beta - \alpha$! Let $\xi : [\alpha, \beta] \to \mathbf{C}^2$ be a continuous zero angular momentum lift of σ to the nonreduced space $\mathbf{C}^2$. (There are a circle's worth of such paths.) The path ξ is a local minimizer of the nonreduced L agrangian action $\int_\alpha^\beta L dt$ among absolutely continuous paths joining $\xi(\alpha)$ to $\xi(\beta)$ in time $\beta - \alpha$. By Marchal's Theorem (see [10] and [4]), $\xi(t)$ is collision-free for $t \in (\alpha, \beta)$ and is a classical solution of Newton's equations. The quotient path $\sigma(t)$ is thus a collision-free solution of the reduced equations with energy $-h$. QED

To prove theorem 3 we will need certain monotonicity properties for the shape potential V. Recall that the shape sphere $\mathbf{CP}^1$ endowed with its Fubini-Study metric (11) is isometric to $(S^2, ds^2/4)$, where ds^2 is the standard metric on the unit sphere S^2. The isometry is

$$St : \mathbf{CP}^1 \to S^2 \subset I\!R^3, \qquad [\xi_1, \xi_2] \mapsto \frac{1}{r^2}(w_1, w_2, w_3)$$

where

$$\begin{aligned} w_1 &= \mu_1|\xi_1|^2 - \mu_2|\xi_2|^2 \\ w_2 + \imath w_3 &= 2\sqrt{\mu_1\mu_2}\xi_2\overline{\xi_1} \\ r^2 &= \mu_1|\xi_1|^2 + \mu_2|\xi_2|^2. \end{aligned} \tag{33}$$

Collinear configurations correspond to the great circle $w_3 = 0$ which we call the collinear equator. The Northern hemisphere ($w_3 \geq 0$) corresponds to positively oriented triangles. The three binary collision points b_{12}, b_{23}, b_{31} lie on the equator and split it into three arcs denoted C_1, C_2, C_3 with the endpoints of C_j being b_{ij} and b_{jk}. Thus C_j consists of collinear shapes in which q_j lies between q_i and q_k.

Introduce standard spherical polar coordinates (ϕ, θ) on the shape sphere so that one of the binary collision points b, say $b = b_{12}$ is the origin ($\phi = 0$) and so that the Northern hemisphere is defined by $0 \leq \theta \leq \pi$. The distance of a point from b relative to the standard metric is ϕ and is a function of r_{12} alone [16]: $\phi = \phi(r_{12}); r_{12} = r_{12}(\phi)$ on the shape sphere. Thus setting $\phi = const.$ is the same as setting $r_{12} = const.$, and defines a geometric circle $\sigma = \sigma(\phi)$ on the shape sphere. The collinear equator is the union of the curves $\theta = 0$ and $\theta = \pi$. The arcs C_1 and C_2 are adjacent to b and we can choose coordinates so that C_1 is contained in the half equator $\theta = 0$ while C_2 contained in the other half equator $\theta = \pi$. In the case of equal masses, $\theta = \pi/2$ defines the isosceles circle through the collision point $b = b_{12}$, lying in the Northern hemisphere.

Lemma 2. *Let σ be any of the half-circles $\phi = const$ centered at b_{12} and lying in the Northern Hemisphere of the shape sphere. Then σ can be parameterized so that $V|_\sigma$, the restriction of V to σ, is strictly convex. Consequently $V|_\sigma$ has a unique minimum. If that minimum is an endpoint of σ then it lies on the 'distant' collinear arc C_3. Otherwise, the minimizer is an interior minimum and coincides with the intersection of σ with the isosceles arc $r_{13} = r_{23}$.*

Remark 1: In all cases, an endpoint of σ which lies on arc C_1 or C_2 is a strict local maximum of the restricted V.

Remark 2: The minimum of $V|_\sigma$ is an endpoint of σ if and only if σ does not intersect the isosceles arc at an interior point.

Remark 3: When $m_1 = m_2$ the isosceles circle in the Northern hemispher (defined by $\theta = \pi/2$) bisects all the circles σ and so the minimum of $V|_\sigma$ always occurs at this bisection point and thus is always interior.

Proof. Use squared length coordinates $s_i = r_{jk}^2$ introduced by Lagrange and re-introduced to us by Albouy [1]. The s_i are subject to the constraints

$$s_i \geq 0$$

and

$$2s_1s_2 + 2s_2s_3 + 2s_3s_1 - s_1^2 - s_2^2 - s_3^2 \geq 0 \tag{34}$$

which together define a convex cone in the 3-space with coordinates s_i. The points of this cone faithfully represent congruence classes of triangles, where two triangles related by reflection are now considered equivalent. The origin of the second inequality is Heron's formula for the signed area Δ of a triangle: $16\Delta^2 = 2s_1s_2 + 2s_2s_3 + 2s_3s_1 - s_1^2 - s_2^2 - s_3^2$.

In squared length coordinates

$$r^2 = \mu_1 s_1 + \mu_2 s_2 + \mu_3 s_3 \tag{35}$$

$$U = m\left(\frac{\mu_1}{\sqrt{s_1}} + \frac{\mu_2}{\sqrt{s_2}} + \frac{\mu_3}{\sqrt{s_3}}\right). \tag{36}$$

A half-circle σ as in the lemma is defined by the two linear constraints $r^2 = 1$, $s_3 = r_{12}^2 = c$ in the s_i. Consequently, such a half-circle is represented by a closed interval whose endpoints correspond to the collinear triangles at which the inequality (34) becomes an equality. U is a strictly convex function of (s_1, s_2, s_3) in the region $s_i \geq 0$ and V on the half-circle becomes $U(s_1, s_2, s_3)$ restricted to this closed interval. Now a strictly convex function restricted to a convex set remains strictly convex, so relative to any affine coordinate on the interval which represents σ we see that V is strictly convex, as claimed.

We can parameterize the half-circle σ by

$$\begin{cases} s_1(u) &= \frac{1}{\mu_1}(1 - \mu_2 u - \mu_3 c) \\ s_2(u) &= u \\ s_3(u) &= c \end{cases}$$

where $u \in [u_-, u_+]$, with $u = u_-$ corresponding to $\theta = 0$, and $u = u_+$ corresponding to $\theta = \pi$. A simple computation gives

$$\frac{d}{du}V(s_1(u), s_2(u), s_3(u)) = \frac{m\mu_2}{2}\left(s_1(u)^{-3/2} - s_2(u)^{-3/2}\right), \tag{37}$$

which proves that $V|_\sigma$ has an interior critical point if and only if σ cross the isosceles curve $s_1 = s_2$. Convexity implies this crossing point is the unique minimum.

If the starting point $u = u_-$ of σ is on C_1, then at this point we have 1 between 2 and 3 so that $r_{12} + r_{13} = r_{23}$. Squaring, we see that $s_2 > s_1$ from which it follows that the derivative (37) is negative at $u = u_-$, and showing that this point is a local maximum for V_σ. A similar argument shows that if an endpoint of σ lies on C_3 then that endpoint is also a local maximum of V_σ. (If the endpoint of σ, corresponding to $u = u_+$ lies on C_3 then the derivative (37) is positive.) QED

We prove now continuity properties of the JM-distance $d_{JM}(q_0, q_1)$ introduced just before the proof of Lemma 1. If q_0 and q_1 are collinear we denote by $d^c_{JM}(q_0, q_1)$ the infimum of the JM-action among collinear rectifiable curves in Q_h joining q_0 to q_1. In a similar way, $d^c_{JM}(q_0, \partial Q_h)$ denotes the minimum of the JM-action among collinear rectifiable curves in Q_h joining q_0 to the Hill boundary.

Lemma 3. *$d_{JM} : Q_h \times Q_h \to I\!R$ and $d^c_{JM} : C_h \times C_h \to I\!R$ are continuous functions.*

Proof. By the triangle inequality it suffices to prove that for every $q_0 \in Q_h$ or $c_0 \in C_h$ the functions $q \to d_{JM}(q_0, q)$ or $q \to d^c_{JM}(c_0, q))$ are continuous in q_0 or in c_0. We give the proof only for $q \to d_{JM}(q_0, q)$. The proof for $q \to d^c_{JM}(q_0, q)$ is similar.

If q_0 is a non-collision shape, not lying on the Hill boundary the $d_{JM}(q_0, q)$ is a regular Riemannian distance in a neighborhood of q_0, so the result is classical. If q_0 lies on the Hill boundary, the function is a pseudo-distance: $d(q_0, q) = 0$ with $q_0 \neq q$ can happen), but the classical result goes through with no changes. So assume now that q_0 is a collision point. We consider two case.

1st case : q_0 is the total collision. Let us choose any non-collision collinear configuration on the shape sphere $\hat{s} \in S^2$. If $q = (r, s) \in Q_h$, and γ is the path obtained by joining the total collision to $\hat{q} = (r, \hat{s})$ along the ray through $\hat{q}$, and then joining $\hat{q}$ to q along a segment of geodesic of $\{r\} \times S^2$, a simple computation

shows the existence of a constant $C > 0$, independent of q, such that $A_{JM}(\gamma) \leq C\sqrt{r} + \mathcal{O}(r\sqrt{r})$, hence $d_{JM}(q_0, q) \to 0$ as $q \to q_0$.

2nd case : q_0 is a partial collision. Assume that q_0 is on the collision ray $r_{12} = 0$. Introduce spherical coordinates (ϕ, θ) as before centered at the binary collision b_{12}, so that $\phi = 0$ corresponds to b_{12}, and $\theta = 0$ or π corresponds to the collinear circle. If r_0 is the radial coordinate of q_0 and $q_1 \in Q_h$ is a point with radial coordinate r_1 and spherical coordinates (ϕ_1, θ_1), let us choose a $\hat{\phi} \in (0, \pi)$ and define the following three paths in spherical coordinates

$$\begin{array}{lll} \gamma_1 : & r(t) = r_0, \quad \phi(t) = t\hat{\phi}, \quad \theta(t) = \theta_1, & t \in [0,1] \\ \gamma_2 : & r(t) = (1-t)r_0 + tr_1, \quad \phi(t) = \hat{\phi}, \quad \theta(t) = \theta_1, & t \in [0,1] \\ \gamma_3 : & r(t) = r_1, \quad \phi(t) = (1-t)\hat{\phi} + t\phi_1, \quad \theta(t) = \theta_1, & t \in [0,1]. \end{array}$$

Let γ be the path obtained by pasting together γ_1, γ_2 and γ_3. This γ joins q_0 to q_1. By (6) and (33) we get

$$V(\phi, \theta) = \frac{c}{\phi} + o(\frac{1}{\phi}),$$

where $c > 0$ depends only on the masses. If $r_0/2 \leq r_1 \leq 3r_0/2$, a simple computation shows that

$$\begin{aligned} A_{JM}(\gamma_1) &= \sqrt{2cr_0\hat{\phi}} + \mathcal{O}(\hat{\phi}^{3/2}), \\ A_{JM}(\gamma_2) &\leq \tfrac{2|r_1 - r_0|}{\sqrt{r_0}} \left(\sqrt{\tfrac{c}{\hat{\phi}}} + o(1/\hat{\phi}) \right), \\ A_{JM}(\gamma_3) &\leq \sqrt{2cr_1\hat{\phi}} + \mathcal{O}(\hat{\phi}^{3/2}), \end{aligned}$$

hence chosing $\hat{\phi} = |r_1 - r_0|$ we get

$$d_{JM}(q_0, q_1) \leq A_{JM}(\gamma) \leq C\sqrt{|r_1 - r_0|} + o(\sqrt{|r_1 - r_0|}),$$

where $C > 0$ depends only on r_0, hence $d_{JM}(q_0, q_1) \to 0$ as $q_1 \to q_0$. QED

Proof of Theorem 3. Let $q_0 \in C_h$ be a collinear configuration. By Proposition 15 there exists a JM-minimizer $\gamma : [0, T] \to Q_h$ among paths starting in q_0 and ending on the Hill boundary ∂Q_h. The JM-action of a path on the Hill boundary is zero, therefore if t_0 is the first time where $\gamma(t_0) \in \partial Q_h$, then $\gamma(t) \in \partial Q_h$ for all $t \in [t_0, T]$. We may cut out the arc $\gamma\big|_{[t_0,T]}$ from γ, or equivalently, assume that $\gamma(t) \notin \partial Q_h$ for all $t < T$. For every $t \in (0, T)$ the restriction $\gamma\big|_{[0,t]}$ is a fixed end point JM-minimizer. By Lemma 1, γ is a classical brake solution, and is collision-free for $t \in (0, T]$.

Let us prove now that if γ is not collinear at all times, it has a unique syzygy : the point q_0. Indeed, assume for the sake of contradiction that $\gamma(\tau)$ is collinear for some $0 < \tau \leq T$. Recall that the set of initial conditions tangent to the collinear submanifold yield collinear curves. Collinear brake points (with zero velocity) are such initial conditions. Consequently, $\tau = T$ is impossible, for if $\gamma(T) \in \partial Q_h$ were collinear, all of γ would be collinear. So $\tau < T$. By the same reasoning, γ cannot be tangent to the collinear manifold at $\tau < T$. Consider the path obtained by keeping $\gamma\big|_{[0,\tau]}$ as it is and reflecting $\gamma\big|_{[\tau,T]}$ with respect to the syzygy plane. This new curve has the same JM-action as γ, and joins q_0 to ∂Q_h, so it is a JM-minimizer. But this new path is not differentiable at $t = \tau$, contradicting the fact that minimizers are solutions, hence smooth. We can now assume without loss of generality that γ is all the time in the half-space of positive oriented shape.

If q_0 is the total collision we will now show that γ is the Lagrange homothety solution. Write $\gamma(t) = (r(t), s(t)) \in Q_h \subset I\!R_+ \times S^2$. Denote by $s_L \in S^2$ the Lagrange shape in the Northern hemisphere. We want to show $s(t) = s_L$ for all t. It is well known that the Lagrange shape is the unique minimum of the shape potential V. Let r_L be the number such that $\frac{1}{r_L}V(s_L) = h$. Suppose that $s(T) \neq s_L$. Then $r(T) > r_L$ since $V(s(T)) = r(T)h > V(s_L) = r_L h$. Since $r(0) = 0$ there will be a smallest number $t = \tau_*$ in the interval $[0, T]$ such that $r(t) = r_L$. Then $r(t) < r_L$ for $t < \tau_*$. The path $\gamma_L(t) = (r(t), s_L)$, $0 \leq t \leq \tau_*$ lies in Q_h, joins q_0 to ∂Q_h and satisfies $U(\gamma_L(t)) \leq U(\gamma(t)$ on this interval (with equality if and only if $s(t) = s_L$). The kinetic energy of this new path is pointwise the same or less than that of the old path, over their common domain. Consequently $A_{JM}(\gamma_L) \leq A_{JM}(\gamma)$, with equality if and only if $s(t) = s_L$ for all $t \in [0, T]$ (and $\tau_* = T$) as desired.

Assume now q_0 is a double collision, say on the collision ray $r_{12} = 0$. We prove that our minimizer γ is not collinear. Assume, for the sake of contradiction that $\gamma(t)$ is collinear at all time $t \in [0, T]$. Since γ is collision free except for q_0, γ is contained in an angular sector of the syzygy plane between the collision ray $r_{12} = 0$ and one of the two other collision rays. Say it lies in the sector C_1 between $r_{12} = 0$ and $r_{13} = 0$. Introduce spherical coordinates (ϕ, θ) on the shape sphere as in Lemma 2. Then $\gamma(t)$ has coordinates $(r(t), \phi(t), \theta(t)) = (r(t), \phi(t), 0)$, where we have arranged the coordinates so that the sector within which γ lies, sector C_1, is given by $\theta = 0$. To re-iterate, at every moment γ's spherical projection lies in the arc C_1 on the collinear equator. By Lemma 2 there exists $\delta \in (0, \pi)$ such that $V(\phi(t), \delta) < V(\phi(t), 0)$ for all $t \in [0, T]$. Let τ_δ be the first time in our interval such that $V(\phi(t), \delta)/r(t) = h$. (There is such a time by an argument similar to the one above when q_0 was triple collision.) Define the new path $\gamma_\delta(t) = (r(t), \phi(t), \delta)$, $t \in [0, \tau_\delta]$. Geometrically, γ_δ is obtained by rotating $\gamma(t)$ by an angle δ around the collision ray $r_{12} = 0$ and keeping only that part of it which stays inside Q_h. Such a rotation is an isometry for the metric (1), but strictly decreases U. Therefore $A_{JM}(\gamma_\delta) < A_{JM}(\gamma)$. But γ is a minimizer, so we have a contradiction.

Using the notations of Lemma 3 we have proved that if q_0 is a collision, we have $d_{JM}(q_0, \partial Q_h) < d^c_{JM}(q_0, \partial Q_h)$.

Assume now that $\overline{q_0} \in C_h$ is a collision configuration. We show the existence of a neighborhood $\mathcal{O}(\overline{q_0})$ of $\overline{q_0}$ in C_h such that for every $q_0 \in \mathcal{O}(\overline{q_0})$, no JM-minimizer from q_0 to ∂Q_h is collinear. Indeed, let us assume, for the sake of contradiction, the existence of a sequence $(q_n)_{n\in\mathbf{N}}$ in C_h, converging to $\overline{q_0}$, with corresponding collinear minimizers. Then $d_{JM}(q_n, \partial Q_h) = d^c_{JM}(q_n, \partial Q_h)$ for all n. By the triangle inequality we have

$$\begin{aligned} |d_{JM}(q_n, \partial Q_h) - d_{JM}(\overline{q_0}, \partial Q_h)| &\leq d_{JM}(q_n, \overline{q_0}), \\ |d^c_{JM}(q_n, \partial Q_h) - d^c_{JM}(\overline{q_0}, \partial Q_h)| &\leq d^c_{JM}(q_n, \overline{q_0}). \end{aligned}$$

Applying Lemma 3 we get $d_{JM}(\overline{q_0}, \partial Q_h) = d^c_{JM}(\overline{q_0}, \partial Q_h)$, which is a contradiction. The neighborhood $\mathcal{U}$ of of the collision locus described in the Theorem can be taken to be the union of the $\mathcal{O}(\overline{q_0})$ as $\overline{q_0}$ varies over the collision locus. This finishes the proof of Theorem 3.

Proof of Theorem 4. In order to prove part (a) of the Theorem, assume $m_1 = m_2$ and let q_0 be a configuration on the collision ray $r_{12} = 0$. Let $\gamma : [0, T] \to Q_h$ be a JM-minimizer from q_0 to ∂Q_h. If q_0 is the triple collision point, by Theorem 3 we know that γ is the Lagrange brake solution, which is an isosceles brake orbit. If q_0 is a double collision point, then express γ in terms of the spherical coordinates

system (r, ϕ, θ) on the shape space S^2, centered at the double collision b_{12} as in Lemma 2 and in the proof of Theorem 3. We can assume that $\gamma(t)$ lies in the Northern hemisphere on the upper half-space $0 \leq \theta \leq \pi$. Because $m_1 = m_2$, the isosceles curve $r_{13} = r_{23}$ in the Northern hemisphere corresonds to the great circle $\{\theta = \pi/2\}$. By Lemma 2, for fixed ϕ, the absolute minimum of $\theta \mapsto V(\phi, \theta)$ is achieved at the isosceles curve $\theta = \pi/2$. Let $0 < \tau \leq T$ be the first time such that $V(\phi(t), \pi/2)/r(t) = h$. Then the path $\tilde{\gamma}(t) = (r(t), \phi(t), \pi/2)$, $t \in [0, \tau]$, joins q_0 to ∂Q_h, has $U(\tilde{\gamma}(t)) \leq U(\gamma(t))$ with equality if and only if $\tilde{\gamma} = \gamma$ on its domain, and has kinetic energy less than or equal to that of γ on its domain. Thus $A_{JM}(\tilde{\gamma}) \leq A_{JM}(\gamma)$, with equality if and only if $\gamma(t)$ is isosceles at all times $t \in [0, T]$. Since γ is a minimizer, it is necessarily isosceles. This proves part (a).

To prove part (b), assume $r_{13} < r_{23}$ at the starting point q_0. The isosceles set $r_{13} = r_{23}$ is invariant by the flow. If we assume, for the sake of contradiction, that $\gamma : [0, T] \to Q_h$ intersects $r_{13} = r_{23}$ at some time $0 < \tau \leq T$, then it is necessarily transverse to $r_{13} = r_{23}$, otherwise $\gamma(t)$ would be isosceles at all times. The reflection with respect to the isosceles plane $r_{13} = r_{23}$ is a symmetry for the kinetc energy and for the potentiel U, therefore the path $\tilde{\gamma}$ obtained by keeping $\gamma\big|_{[0,\tau]}$ as it is, and reflecting $\gamma\big|_{[\tau,T]}$ about the isosceles plane $r_{13} = r_{23}$, is still a JM-minimizer. But $\tilde{\gamma}$ is not differentiable at $t = \tau$. This is a contradiction.

The proof of part (c) is obtained applying (b) to the three isosceles plane $r_{ij} = r_{ik}$. QED.

Proof of Proposition 14: Seifert's Tubular Neighborhood Theorem.

Our construction is the same as Seifert's [22]. The one real difference is that our Hill boundary ∂Q_h is not compact, whereas his is compact.

We will construct Seifert's neighborhood in the non-reduced configuration space $\mathbf{C}^2$ and then project it via the quotient map π of eq (10) to Q_h to arrive at the reduced Seifert neighborhood. We abuse notation by using the same symbol Q_h for the reduced and non-reduced Hill regions and similarly using the same symbol $S_h^\epsilon \subset Q_h$ for the reduced and non-reduced Seifert neighborhoods to be constructed.

The configuration space $\mathbf{C}^2$ is endowed with the mass inner product $<,>$ which is the real part of the Hermitian inner product of eq. (7). The equations there are

$$\ddot{q} = \nabla U(q), \tag{38}$$

where the gradient of U is calculated with respect to the mass inner product. The Jacobi-Maupertuis action of a curve $\gamma : [t_0, t_1] \to Q_h$ is defined by

$$A_{JM}(\gamma) = \int_{t_0}^{t_1} \sqrt{K}\sqrt{2(U-h)}dt,$$

where $K = \|\dot{\gamma}\|^2/2$.

As a first step we construct a map $F : \partial Q_h \times [0, \overline{\delta}] \to Q_h$ which is an analytic diffeomorphism onto its image and is such that the curves $t \mapsto F(x, t^2)$ are the brake solutions starting from x at $t = 0$. Given $\alpha < h < \beta$, define the open set

$$D_{\alpha,\beta} = \{x \in \mathbf{C}^2, \quad \alpha < U(x) < \beta\}.$$

If $x \in D_{\alpha,\beta}$, then the smallest of the mutual distances r_{ij} of the triangle defined by x is bounded below by a constant depending only on β and the masses. We denote by $q_x(t)$ the solution to (38) with initial conditions $q_x(0) = x$ and $\dot{q}_x(0) = 0$. (Note: these solutions need not have energy h.) Choose positive numbers a and b

such that $a < \alpha < h < \beta < b$. By classical results on differential equations, there exists $T > 0$ such that for every $x \in D_{\alpha,\beta}$, the solution $q_x(t)$ is well defined for $t \in [-T, T]$ and satisfies $q_x(t) \in D_{a,b}$. The map $(x, t) \mapsto q_x(t)$ is analytic, even in t (i.e. $q_x(-t) = q_x(t)$), and its derivatives up to the second order are uniformly bounded on $D_{\alpha,\beta} \times [-T, T]$. By equations of motion (38)

$$q_x(t) = x + \frac{\nabla U(x)}{2} t^2 + \mathcal{O}(t^4).$$

Let us set $T_1 = \sqrt{T}$. Since $q_x(t)$ is even in t, the map $F(x, \tau) = q_x(\sqrt{\tau})$, $(x, \tau) \in D_{\alpha,\beta} \times [0, T_1]$ is still analytic, so it can be extended to negative values of τ, giving an analytic map $F : D_{\alpha,\beta} \times [-T_1, T_1] \to D_{a,b}$ satisfying

$$F(x, \tau) = x + \frac{\nabla U(x)}{2} \tau + f(x, \tau), \tag{39}$$

where the $f(x, \tau) = \mathcal{O}(\tau^2)$, uniformly for $x \in D_{\alpha,\beta}$.

Define

$$G : D_{\alpha,\beta} \times [-T_1, T_1] \to D_{a,b} \times I\!R, \qquad G(x, \tau) = (F(x, \tau), U(x)).$$

By (39), the differential of G at a point $(x, \tau = 0)$ gives

$$DG_{(x,0)}(\delta x, \delta\tau) = (\delta x + \frac{\nabla U(x)}{2} \delta\tau, < \nabla U(x), \delta x >).$$

An easy computation shows that $DG_{(x,0)}$ is invertible and $\|(DG_{(x,0)})^{-1}\|$ is bounded as $x \in D_{\alpha,\beta}$. Moreover, since all derivatives of G up to the second order are uniformly bounded on $D_{\alpha,\beta} \times [-T_1, T_1]$, we can apply a strong version of the inverse function theorem and find a $\overline{\delta} > 0$ such that for all $x \in D_{\alpha,\beta}$, the map G defines a diffeomorphism from $B(x, \overline{\delta}) \times [-\overline{\delta}, \overline{\delta}]$ into its image, where $B(x, \overline{\delta})$ denotes the closed ball centered in x with radius $\overline{\delta}$. If we take the restriction to $U(x) = h$ and define $B_{\partial Q_h}(x, \overline{\delta}) = B(x, \overline{\delta}) \cap \partial Q_h$ we find that F defines an analytic diffeomorphism from $B_{\partial Q_h}(x, \overline{\delta}) \times [-\overline{\delta}, \overline{\delta}]$ into its image. Let us prove now that by decreasing sufficiently $\overline{\delta}$, the restriction of F to $\partial Q_h \times [-\overline{\delta}, \overline{\delta}]$ is an analytic diffeomorphism onto its image. Indeed, we have proven that at every point it is a local diffeomorphism. Assume, for the sake of contradiction, there exist two sequence $(x_n, \tau_n)_{n\in\mathbf{N}}$ and $(x'_n, \tau'_n)_{n\in\mathbf{N}}$ satisfying $(x_n, \tau_n) \neq (x'_n, \tau'_n)$ such that $\tau_n \to 0$, $\tau'_n \to 0$ as $n \to +\infty$ and $F(x_n, \tau_n) = F(x'_n, \tau'_n)$ for all $n \in \mathbf{N}$. By (39) and uniform boundedness of ∇U on ∂Q_h, we have $\|x_n - x'_n\| \to 0$ as $n \to +\infty$, therefore $x'_n \in B_{\partial Q_h}(x_n, \overline{\delta})$ if n is sufficiently great. This contradicts the fact that F is a diffeomorphism on every $B_{\partial Q_h}(x, \overline{\delta}) \times [-\overline{\delta}, \overline{\delta}]$.

The second step consists in defining a diffeomorphism Φ on $\partial Q_h \times [0, \overline{\delta}]$ of the form $\Phi(x, y) = F(x, \beta(x, y))$ where the scalar function β allows us to "straighten out" the JM action. Let $S(x, \tau)$ denotes the JM-length of the extremal $F(x, s)_{s\in[0,\tau]}$. From (39) we compute

$$S(x, \tau) = \tau^{3/2} \left(\frac{1}{3} \|\nabla U(x)\|^2 + g(x, \tau) \right), \tag{40}$$

where $g(x, \tau)$ is analytic, and $g(x, \tau) = \mathcal{O}(\tau)$, uniformly on $x \in \partial Q_h$. The map $H(x, \tau) = (x, S(x, \tau)^{2/3})$ is analytic in $\partial Q_h \times [0, \overline{\delta}]$, and we can find $\epsilon > 0$ such that H maps a neighborhood of $\partial Q_h \times \{0\}$ (in $\partial Q_h \times [0, \overline{\delta}]$) diffeomorphically into $\partial Q_h \times [0, \epsilon]$. Then H^{-1} has the form $(x, y) \mapsto (x, \beta(x, y))$. We set $\Phi = F \circ H^{-1}$

where the domain of H^{-1} is $\partial Q_h \times [0,\epsilon]$. It is clear now that for every $x \in \partial Q_h$, and for every $\delta \in (0,\epsilon)$, the curve $(\Phi(x,y))_{y\in[0,\delta]}$ is a segment of brake orbit and its JM-action is $\delta^{3/2}$.

The Jacobi metric g_h in $Q_h \subset \mathbf{C}^2$ is given by

$$g_h(x)(v,v) = 2(U(x)-h)\langle v, v\rangle$$

We now show that brake solutions $(\Phi(x,y))_{y\in[0,\epsilon]}$ are g_h-orthogonal to hypersurfaces $\Phi(\partial Q_h \times \{y\})$. The argument is the standard one used to prove the Gauss lemma in Riemannian geometry. We are to show that $g_h(\partial_y\Phi(x,y), \partial_x\Phi(x,y)\xi) = 0$ for all ξ tangent to the hypersurface. For this purpose, set $f(y;x) = g_h(\partial_y\Phi(x,y), \partial_x\Phi(x,y)\xi)$ for fixed ξ. Because the lengths of the curves $y \to \Phi(x,y)$ up to the value y_0 are all $y_0^{3/2}$ we have that $g_h(\partial_y\Phi(x,y), \partial_y\Phi(x,y)) = \frac{9}{4}y$, independent of x. Differentiating this identity in the ξ-direction tangent to the hypersurface and commuting derivatives we get $g_h(\partial_y\Phi(x,y), \nabla_y\partial_x\Phi(x,y)\xi) = 0$ where ∇ denotes the g_h-Levi-Civita connection. Now, because the y-curves are reparameterized geodesics with lengths only depending on y we have that $\nabla_y\partial_y\Phi(x,y) = \mu(y)\partial_y\Phi(x,y)$ for some function $\mu(y)$. Differentiate $f(y;x)$ with respect to y, and use metric compatibility and the previous equations to derive the linear differential equation $\partial_y f = \mu(y) f(y;x)$ with x as a parameter. But $f(0;x) = 0$ so the initial condition of this differential equation is zero, from which it follows that f is identically 0. (Indeed μ is the second derivative of $y^{3/2}$ with respect to y, which is singular at $y = 0$, but the analysis goes through.)

An arbitrary tangent vector $v \in T_qQ_h$ at $q = \Phi(x,y)$ can be written

$$v = \partial_x\Phi(x,y)\xi + \partial_y\Phi(x,y)\lambda,$$

where $(\xi,\lambda) \in T_x\partial Q_h \times I\!R$. By the previous orthogonality discussion and the fact that $y^{3/2}$ is the arclength of $(\Phi(x,y))_{y\in[0,\epsilon]}$ we have for such v:

$$g_h(v,v) = (U-h)\kappa_h(x,y)(\xi,\xi) + \frac{9y}{4}\lambda^2. \tag{41}$$

where $\kappa_h(x,y)$ is a positive definite quadratic form on $T_x\partial Q_h$. If $\gamma : [t_0,t_1] \to Q_h$ is any rectifiable curve joining a point $q = \Phi(x,\delta)$ (with $0 < \delta < \epsilon$) to ∂Q_h, and if I is the set of times $t \in [t_0,t_1]$ such that $\gamma(t) \in S_h^\epsilon = \Phi(\partial Q_h, [0,\epsilon])$, by (41) we have

$$A_{JM}(\gamma) = \int_{t_0}^{t_1} \sqrt{g_h(\dot\gamma(t),\dot\gamma(t))}dt \geq \int_I \frac{3\sqrt{y(t)}}{2}|\dot y(t)|dt \geq \delta^{3/2}, \tag{42}$$

where we term $(x(t),y(t)) = \Phi^{-1}(\gamma(t))$ for $t \in I$. Moreover, we have equality in (42) if and only if, up to an arc contained in ∂Q_h, the curve γ is a reparametrization of $(\Phi(x,y))_{y\in[0,\delta]}$. The Euclidean length of $(\Phi(x,y))_{y\in[0,\delta]}$ is given by the integral $\int_0^\delta \|\partial_y\Phi(x,y)dy\|\,dy$. By (39), (40) and by definition of Φ we have

$$\Phi(x,y) = x + \frac{\nabla U(x)}{2\cdot 3^{2/3}\|\nabla U(x)\|^{4/3}}y + \mathcal{O}(y^2). \tag{43}$$

Since $\nabla U(x)$ is uniformly bounded on ∂Q_h, there exists a strictly positive constant M, independent of x and δ as long as $\delta < \epsilon$, such that

$$\int_0^\delta \|\partial_y\Phi(x,y)dy\|\,dy \leq M\delta.$$

Let us show now there exists $\alpha > 0$ such that for every $\delta \in (0, \epsilon)$ we have $U \geq h + \alpha\delta$ on $Q_h \setminus S_h^\delta$. By (43) we have

$$U(\Phi(x,y)) = h + \frac{\|\nabla U(x)\|^{2/3}}{2 \cdot 3^{2/3}} y + \mathcal{O}(y^2), \tag{44}$$

for $(x, y) \in \partial Q_h \times [0, \epsilon]$. Since $\|\nabla U(x)\|$ is bounded and uniformly bounded below below by a positive constant, we can find two constant $0 < \alpha < \beta$ such that

$$h + \alpha y \leq U(\Phi(x,y)) \leq h + \beta y$$

for every $(x, y) \in \partial Q_h \times [0, \epsilon]$. Assume now, for the sake of contradiction, the exisence of a point $q \in Q_h \setminus S_h^\delta$ such that $U(q) < h + \alpha\delta$. Since level set of U are connected by arc, there would exist a point q' on $\Phi(\partial Q_h, \delta)$ such that $U(q') < h + \alpha\delta$, and this is in contradiction with (44).

5. Periodic Brake Orbits

The goal of this section is to prove Theorem 2 about the existence of simple, periodic brake orbits for the isosceles three-body problem. Other examples of periodic, isosceles brake orbits, more complicated than the one described here, are given in [25].

Let p be a brake initial condition other than the Lagrange homothetic one, p_0. According the results of section 3.1, p can be followed forward to meet the syzygy submanifold, $\tilde{C}_h$, i.e., the submanifold of the energy manifold with collinear shapes. There is a natural reflection symmetry of the energy manifold through $\tilde{C}_h$, obtained by reflecting the shape variables (x, y) and their velocities (x', y') while leaving the size variables (r, v) unchanged. Call this reflection map R. Then R is a symmetry of the differential equation if one also reverses time. If the orbit, $\gamma(t)$, of p meets $\tilde{C}_h$ orthogonally after time T_1, then reflecting the orbit segment and reversing time gives the continuation of γ to the time interval $[T_1, 2T_1]$ and $\gamma(2T) = R(p)$, the reflection of p. Now it follows from symmetry during the interval $[2T_1, 4T_1]$ the orbit retraces its path, returning to p. So p determines a periodic brake orbit of period $4T$.

If $\gamma(t)$ does not meet $\tilde{C}_h$ orthogonally and if $\gamma(T)$ is not in the local stable manifold of the Lagrange triple collision, then it can be followed to a *second syzygy*, say at time $T_2 > T_1$. If this crossing is orthogonal one obtains a periodic brake orbit of period $4T_2$ by reflection.

We will see that such a second-syzygy periodic brake orbit exists in the isosceles subsystem of the three-body problem, at least for certain choices of the mass parameters. It is not known whether a first-syzygy brake orbit exists. Numerical experiments suggest that no such orbits exist in the equal mass case.

5.1. **The Isosceles Three-Body Problem.** Assume that two masses are equal, say $m_1 = m_2 = 1$. Then there is an invariant submanifold of the three body problem such that the shape remains an isosceles triangle with m_3 on the symmetry axis for all time. Up to rotation, this isosceles subsystem is obtained by making the Jacobi variables $\xi_1, \dot{\xi}_1$ real and $\xi_2, \dot{\xi}_2$ imaginary.

As in section 2 separate size and shape variables will be used. In addition, double collisions will have to be explicitly regularized. A convenient way to do this is to define an angular variable which gives a multiple cover of the isosceles shape space and which is locally a branched double cover near the binary collisions, as in the

familiar Levi-Civita regularization. Using the projective Jacobi variables $[\xi_1, \xi_2]$, this parametrization of the isosceles shapes can be accomplished by setting

$$\xi_1 = \frac{1}{\sqrt{\mu_1}} \cos^2(\theta) \qquad \xi_2 = \frac{i}{\sqrt{\mu_2}} 2 \sin(\theta).$$

The isosceles binary collision shape corresponds to $\xi_1 = 0$ or $\theta = \pm\frac{\pi}{2} \bmod 2\pi$.

Note that with the assumptions about the masses, one has

$$\mu_1 = \frac{1}{2} \qquad \mu_2 = \frac{2m_3}{2 + m_3}.$$

Substitution gives a reduced Lagrangian

$$L_{red}(r, \dot{r}, \theta, \dot{\theta}) = K_0 + \frac{1}{r} V(\theta)$$

where

$$K_0 = \frac{\dot{r}^2}{2} + \frac{2r^2 \cos^2\theta \, \dot{\theta}^2}{(1 + \sin^2\theta)^2}$$

and

$$V(\theta) = (1 + \sin^2\theta) \left(\frac{1}{\sqrt{2} \cos^2\theta} + \frac{2\sqrt{2}\, m_3}{\sqrt{(1 + \sin^2\theta)^2 + \frac{8}{m_3} \sin^2\theta}} \right).$$

Introducing a change of time scale $' = r^{\frac{3}{2}} \cos^2\theta \, \dot{}$ leads to the following system of differential equations:

$$\begin{aligned} r' &= vr \cos^2\theta \\ v' &= \tfrac{1}{2} v^2 \cos^2\theta + \tfrac{1}{4} w^2 (1 + \sin^2\theta)^2 - W(\theta) \\ \theta' &= \tfrac{1}{4} w (1 + \sin^2\theta)^2 \\ w' &= W'(\theta) - \tfrac{1}{2} vw \cos^2\theta + \sin\theta \cos\theta \left(2r + v^2 - \tfrac{1}{2} w^2 (1 + \sin^2\theta) \right) \end{aligned} \tag{45}$$

where $W(\theta) = \cos^2\theta \, V(\theta)$. The energy conservation equation is

$$\tfrac{1}{2} v^2 \cos^2\theta + \tfrac{1}{8} w^2 (1 + \sin^2\theta)^2 - W(\theta) = -r \cos^2\theta \tag{46}$$

where the energy has been fixed at -1.

Since $W(\theta)$ is an analytic function, the binary collisions have been regularized. Moreover the triple collision singularity has been blown up into an invariant manifold at $\{r = 0\}$ as before. The isosceles Hill's region for energy -1 is $Q_1 = \{(r, \theta) : 0 \le r \le V(\theta)\}$. This is shown in figure 6 for the equal mass case $m_3 = 1$.

Using these coordinates, syzygies occur at the Euler shape ($\theta = 0 \bmod 2\pi$) and at binary collision ($\theta = \pm\frac{\pi}{2} \bmod 2\pi$). These are indicated by the bold vertical lines in figure 6. The reflection through syzygy amounts to reflecting the position variables (r, θ) through these vertical lines while taking the velocity (v, w) to $(v, -w)$ and reversing time. To get a symmetric periodic orbit, one needs to reach syzygy with $v = 0$. Since $r' = vr \cos^2\theta$ this is equivalent to orthogonality at the Euler vertical lines $\theta = 0 \bmod 2\pi$, but every orbit crosses the lines $\theta = \pm\frac{\pi}{2} \bmod 2\pi$ orthogonally. A numerically computed periodic brake orbit is shown in the figure. It begins on the zero velocity curve, crosses the syzygy line at $\theta = -\frac{\pi}{2}$ with $v < 0$ (not apparent from the figure), then continues to its second syzygy at $\theta = 0$ where it crosses orthogonally with $v = 0$. The rest of the orbit is obtained by symmetry. The goal

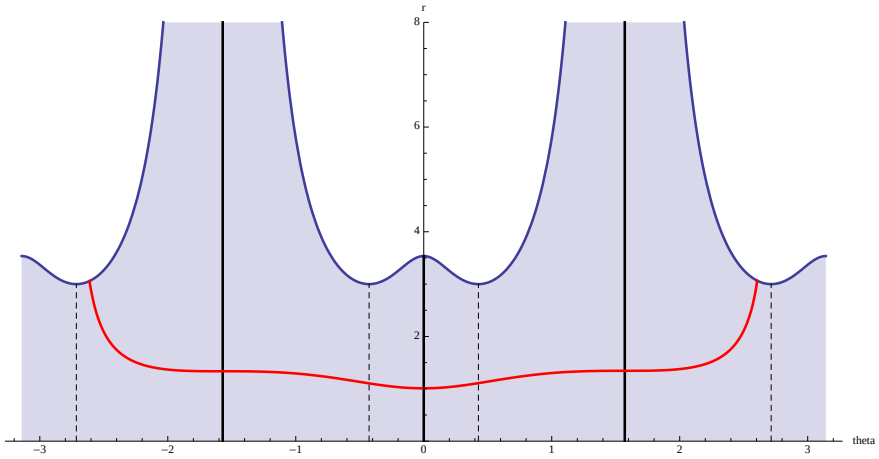

FIGURE 6. Isosceles Hill's region for the equal mass three-body problem using coordinates (θ, r). The zero velocity curve is the top boundary curve. The syzygy configurations are represented by the thick vertical lines ($\theta = 0$ is the collinear central configuration and $\theta = \pm\frac{\pi}{2}$ are binary collision shapes). A numerically computed periodic brake orbit is also shown.

of this section is to prove the existence of such an orbit for certain choices of the parameter m_3.

The idea of the proof can be described briefly as follows. Consider a curve, Z, of brake initial conditions whose shapes vary from equilateral to binary collision. In figure 6, the equilateral shapes (minima of the shape potential) are shown with dashed vertical lines. Let θ^* be the equiliateral shape in the interval $[0, \frac{\pi}{2}]$. Other equilateral shape occur at the points $\pm\theta^* + k\pi$ where k is an integer. The curve Z will consist of the brake initial conditions with $\theta \in [\theta^* - \pi, -\frac{\pi}{2}]$ (in figure 6, this is the part of the top boundary curve between the left-most dashed and left-most bold vertical lines). It will be shown that as the initial condition $p \in Z$ varies, there is at least one point which can be followed to meet the vertical line $\theta = 0$ orthogonally.

The proof will use a geometrical argument in the three-dimensional energy manifold, $P_1 = \{(r, \theta, v, w) : r \geq 0, H = -1\}$. This manifold can be visualized through its projection to (r, v, θ)-space, which is given by the inequality $r + \frac{1}{2}v^2 \leq V(\theta)$. This projection is shown in figure 7. On the top surface of the projection, $w = 0$. The full energy manifold can be viewed as two copies of this projection (one with $w \geq 0$ and one with $w \leq 0$) glued together along this top surface. The desired orbit has the property that the shape angle $\theta(t)$ will increase monotonically from $\theta(0)$ to $\theta(T_2) = 0$ where T_2 denotes the second-syzygy time described above. Therefore it suffices to consider the part of the energy manifold with $\theta' = \frac{1}{4}w(1 + \sin^2\theta)^2 \geq 0$. On this half of the energy manifold, one can solve (46) uniquely for $w(r, v, \theta)$.

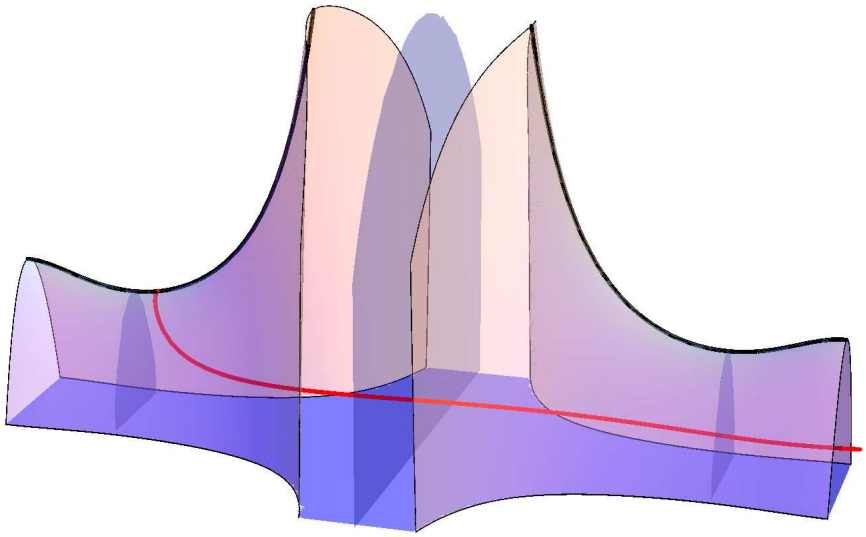

FIGURE 7. Projection of the $w \geq 0$ part of the energy manifold in coordinates θ (width), v (depth), r (height), The top surface is $\{w = 0\}$; the floor is part of the collision manifold $\{r = 0\}$. A numerically computed periodic brake orbit is also shown, passing through the walls of the regions R_I, R_{II}, R_{III} as described in the proof.

The relevant part of the energy manifold will be divided into three regions:

$$R_I = P_1 \cap \{\theta \in [\theta^* - \pi, -\tfrac{\pi}{2}], w \geq 0\}$$
$$R_{II} = P_1 \cap \{\theta \in [-\tfrac{\pi}{2}, -\theta^*], w \geq 0\}$$
$$R_{III} = P_1 \cap \{\theta \in [-\theta^*, 0], w \geq 0\}.$$

The planes $\theta = \theta^* - \pi$ and $\theta = -\frac{\pi}{2}$ will be called the *left and right walls* of R_I respectively with similar definitions for the other regions (the walls are the four vertical planes in figure 7).

These regions contain certain restpoints on the collision manifold which will now be described. Let $L_\pm$ be the Lagrange restpoints in R_I at $(r, \theta, v, w) = (0, \theta^* - \pi, \pm v^*, 0)$ where $v^* = \sqrt{2V(\theta^*)}$. These are connected by the Lagrange homothetic orbit which is the curve of intersection of the left wall of R_I with the boundary surface $\{w = 0\}$. Similarly, there are Lagrange restpoints $L'_\pm$ at $(r, \theta, v, w) = (0, -\theta^*, \pm v^*, 0)$ and a corresponding homothetic orbit in the left wall of R_{III} which is also the right wall of R_{II}. Finally there are Eulerian restpoints $E_\pm$ at $(r, \theta, v, w) = (0, 0, \pm\sqrt{2V(0)}, 0)$ and an Eulerian homothetic orbit in the right wall of R_{III}.

Let Z be the part of the zero velocity curve in R_I. The proof will follow a subset of Z forward under the flow through these three regions to obtain a curve in the right wall of R_{III}, i.e., in the syzygy set at $\theta = 0$. It will be shown that this final curve crosses the plane $v = 0$ and the point of crossing will determine the required periodic brake orbit. The next lemma shows how the flow can be used to carry orbits across the various regions.

Lemma 4. *Regions R_I and R_{III} are positively invariant sets for the flow while R_{II} is negatively invariant. With the exception of the Lagrange homothetic orbits, orbits cross these regions as follows: any orbit beginning in the left wall of region R_I, crosses the region and exits at the right wall. The same hold for R_{III} except for orbits in the stable manifold of E_+ (which is contained in the triple collision manifold $\{r = 0\}$).*

Similarly, except for the Lagrange orbit, any backward-time orbit beginning in the right wall of R_{II} can be followed back to the left wall. Finally, forward orbits beginning in the left wall of R_{II} either leave R_{II} through the right wall, leave R_{II} through the top surface $\{w = 0\}$ or converge to one of the Lagrange restpoints $L'_\pm$ in the right wall as $s \to \infty$.

Proof. By definition, an orbit in any of the three regions satisfies $\theta' \geq 0$ so $\theta(s)$ is non-decreasing. Referring to figure 7, the lower boundary surface $\{r = 0\}$ is invariant. On the upper boundary surface, $w = 0$ and

$$w' = W'(\theta) + \sin\theta\cos\theta\left(2r + v^2\right) = W'(t) + 2\sin\theta\cos\theta W(\theta) = \cos^2\theta\, V'(\theta).$$

Now R_I, R_{III} are are defined by θ-intervals where $V'(\theta) \geq 0$ so for these regions, $w' \geq 0$ on the top boundary. This proves positive invariance and also that it is only possible to leave through the right wall. Similarly, in R_{II} we have $w' \leq 0$ on the top wall so the region is negatively invariant and backward orbits can only leave through the left wall.

Assume for the sake of contradiction, that an orbit other than the Lagrange homothetic orbit remains in R_I for all time $s \geq 0$. Then $\theta(s)$ converges monotonically to some limit $\theta_\infty \in (\theta^* - \pi, -\frac{\pi}{2}]$ (it cannot be $\theta^* - \pi$ since all points in this plane but not on the Lagrange orbit have $w > 0$ so they initially move to the right). The omega limit set is either empty or else it must be a nonempty invariant subset of $\{\theta = \theta_\infty\}$. However, these planes don't contain any nontrivial invariant sets. So the omega limit set must be empty which is only possible if $\theta_\infty = -\frac{\pi}{2}$ and if the orbit leaves every compact subset of the energy manifold. To show that this is impossible, let $\lambda = \sqrt{2r + v^2}$ so that the energy equation becomes

$$\tfrac{1}{2}\lambda^2\cos^2\theta + \tfrac{1}{8}w^2(1 + \sin^2\theta)^2 = W(\theta).$$

We will show that $\lambda(s)$ remains bounded to complete the argument. From (45) we find $\lambda\lambda' = \frac{1}{8}vw^2(1 + \sin^2\theta)^2$. Since $\theta(s)$ is increasing, we can reparametrize by θ to get $\frac{d\lambda}{d\theta} = \frac{1}{2}\frac{vw}{\lambda}$ and since $|v| \leq \lambda$ we have

$$\left|\frac{d\lambda}{d\theta}\right| \leq \tfrac{1}{2}w. \tag{47}$$

As $w(s)$ is bounded by the energy relation, we get a bound for $|\frac{d\lambda}{d\theta}|$. It follows that $\lambda(s)$ is bounded along the part of the orbit in R_I as required.

The same argument applies to backward-time orbits in R_{II}, showing that they must reach the right wall. The argument for forward orbits in R_{III} is easier since

this region is compact. The omega limit set of an orbit remaining in R_{III} for all time would have to be a nonempty invariant set in a plane $\theta = \theta_\infty \in (-\frac{\pi}{2}, 0]$. The only invariant sets are the Eulerian restpoints and homothetic orbit and so the omega limit set would have to be one of the restpoints. However, the stable manifold of E_- is just the restpoint itself and the homothetic orbit, so the orbit must be in the stable manifold of E_+ which, it so happens, is contained in the collision manifold. QED

In addition to this lemma about region-crossing, it will be necessary to use some facts about the flow on the isosceles triple collision manifold. Setting $r = 0$ in (45) gives the dynamics on the triple collision manifold. The energy equation (46) gives

$$\tfrac{1}{2}v^2 \cos^2\theta + \tfrac{1}{8}w^2(1+\sin^2\theta)^2 - W(\theta) = 0.$$

Using this to eliminate w gives a flow on part of the (θ, v) plane satisfying $v^2 \leq 2V(\theta)$. This is shown in figure 8 for the case $m_3 = 1$. An important property of the flow is that it is gradient-like with respect to the variable v. Indeed, using the energy equation gives $v' = \frac{1}{8}w^2(1+\sin^2\theta)^2 \geq 0$ and it can be shown that $v(s)$ is strictly increasing except at the restpoints. The restpoints $L_\pm, L'_\pm$ are saddle points.

Certain properties of their stable and unstable manifolds will be used in the proof. Let γ, γ' denote the branches of $W^u(L_-), W^u(L'_-)$ in $\{w > 0\}$ (bold lines in figure 8). The key properties needed to complete the existence proof refer to the intersections of these branches with the syzygy lines at $\theta = -\frac{\pi}{2}$ and $\theta = 0$. We require that γ remains in $\{w > 0\}$ at least until it crosses these two syzygy lines and that the intersection points are $(\theta, v) = (-\frac{\pi}{2}, v_1)$ where $v_1 < 0$ and $(\theta, v) = (0, v_2)$ with $v_2 > 0$. Furthermore we require that γ' remains in $\{w > 0\}$ at least until it crosses $\{\theta = 0\}$ at a point $(\theta, v) = (0, v_3)$ with $v_3 < 0$. Call a mass parameter m_3 *admissible* if these hypotheses hold. The figure indicates that $m_3 = 1$ is admissible and the next lemma guarantees that this is so. The proof will be given later.

Lemma 5. *There is a nonempty open set of admissible masses which including the equal mass value $m_3 = 1$.*

These lemmas can be applied to follow a subset of Z through R_I to a curve Z_I in the right wall. The left endpoint of Z is the brake initial condition with equilateral shape, $\theta = \pi - \theta^*$. Call this point p_0. The corresponding orbit is the Lagrange homothetic orbit, so it remains in R_I for all $t \geq 0$ and converges to the restpoint L_-. It follows from lemma 4 that $Z \setminus p_0$ can be followed through R_I to the right wall. It will be important to understand the image curve Z_I whose three-dimensional projection lies in the half-plane $\{(r, \theta, v) : r \geq 0, \theta = -\frac{\pi}{2}\}$. Use (v, r) as coordinates in this half-plane and let p_I be the point $(v, r) = (v_1, 0)$ where the branch γ of $W^u(L_-)$ meets the right wall.

Lemma 6. *Assume m_3 is admissible. Then the image curve Z_I of $Z \setminus p_0$ is a continuous open arc in the right wall of R_I. One end converges to $p_I = (v_1, 0)$ and at the other end $|(r, v)| \to \infty$.*

Proof. The initial curve $Z \setminus p_0$ is a continuous open arc in the zero velocity curve from p_0 to $(r, \theta, v, w) = (\infty, -\frac{\pi}{2}, 0, 0)$. The flow across the right wall of R_I is transverse since $w > 0$ there. So Z_I is a continuous open arc. Initial conditions near p_0 on Z will follow the Lagrange homothetic orbit to a neighborhood of the

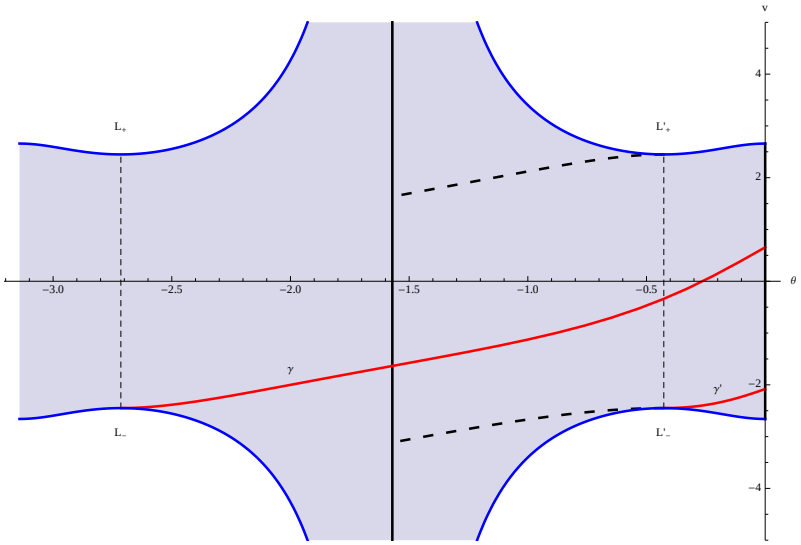

FIGURE 8. Flow on the $w > 0$ part of the collision manifold in the equal mass case (the "floor" in figure 7). The crucial branches γ, γ' of $W^u(L_-), W^u(L'_-)$ are shown. γ intersects the line $\theta = -\frac{\pi}{2}$ with $v < 0$ and $\theta = 0$ with $v > 0$ and γ' intersects $\theta = 0$ with $v < 0$. Two shorter branches of stable manifolds which play a role in the proof are also shown.

restpoint L_- and then follow the branch γ of the unstable manifold to meet the right wall near the point p_I.

Next consider an initial point near the other end of the curve. Initially the quantity $\lambda = \sqrt{2r + v^2}$ is large and $\theta \approx -\frac{\pi}{2}$. It follows from (47) that λ is still large when the orbit meets the right wall. QED

Now part of the curve Z_I will be followed across region R_{II}. Unfortunately, this region is not positively invariant (solutions can leave by w becoming negative, i.e., θ begins to decrease). It turns out, however, that part of the curve Z_I is trapped inside R_{II} by an invariant surface, namely, the stable manifold $W^s(L'_-)$. From the linearization at the restpoint, it follows that $W^s(L'_-)$ has dimension 2. One of the orbits in the stable manifold is the Lagrange homothetic orbit which lies in the right wall of R_{II} connectin L'_+ to L'_-. It is known that $W^u(L'_+)$ and $W^s(L'_-)$ intersect transversely along this orbit [24]. $W^s(L'_-)$ also contains two orbits in the collision manifold, one of which lies in region R_{II}. The one-dimensional manifold $W^s(L'_+)$ also contains an orbit in R_{II}. These two branches of stable manifolds are shown as dashed curves in figure 8.

Consider the "quadrant" of the surface $W^s(L'_-)$ in region R_{II}. One edge is the orbit in the collision manifold just described and the other is the Lagrange homothetic orbit in the right wall. It follows from lemma 4 that with the exception of the homothetic orbit itself, orbits in this quadrant can be followed backward under the flow to reach the left wall of the region. The intersection of the surface and the wall will be a curve. One endpoint of the curve arises from the branch of $W^s(L'_-)$ in the collision manifold. Since v is decreasing for backward orbits

in the collision manifold, this endpoint will be of the form $(v, r) = (v_0, 0)$ with $v_0 < -v^* < v_1 < 0$. To find the other endpoint, note that backward orbits in the quadrant near the homothetic orbit will follow the homothetic orbit back near the restpoint L'_+ and then follow the branch of $W^s(L'_+)$ in R_{II}. This branch of stable manifold is related by symmetry to the branch γ of $W^u(L_-)$. Hence it meets the left wall of R_{II} at $(v, r) = (-v_1, 0)$, $-v_1 > 0$.

So the surface $W^s(L'_-)$ intersects the left wall of R_{II} in a curve connecting the two points $(v_0, 0), (-v_1, 0)$ in the collision manifold but otherwise lying in $\{r > 0\}$. It follows from lemma 6 that Z_I crosses this curve. Let Z'_I denote the part of Z_I below the stable manifold from $(v, r) = (-v_1, 0)$ to the first intersection, call it q_I, with $W^s(L'_-)$. Points in $Z'_I \setminus q_I$ can be followed forward across region R_{II} to its right wall.

Indeed if $p \in Z'_I \setminus q_I$, the forward orbit of p remains in the part of R_{II} below the invariant manifold $W^s(L'_-)$. By lemma 4, it either reaches the right wall or converges to $L'_\pm$. By construction, the only point of Z'_I in $W^s(L'_\pm)$ is q_I. Points of Z'_I near q_I will reach the right wall of R_{II} near L'_-. The other endpoint of Z'_I is the point p_I with $(v, r) = (v_1, 0)$ in the branch of $W^u(L_-)$ described in the definition of admissible mass. This point continues to follow that branch to its intersection point with the right wall. Thus the image of $Z'_I \setminus q_i$ is an open arc Z_{II} in the right wall of R_{II} connecting L'_- to γ and otherwise lying in $\{r > 0\}$.

Since Z_{II} does not intersect the Lagrange homothetic orbit or the stable manifold of E_+, lemma 4 shows that it can be followed across region R_{III} to the right wall to form an open arc Z_{III}. The endpoints of Z_{III} are at the point $(v, r) = (v_2, 0)$ and $(v_3, 0)$ where the branches of unstable manifolds γ and γ' cross the wall. Since m_3 is admissible, we have $v_3 < 0 < v_2$, so the is at least one point of Z_{III} with $v = 0$. Thus there is a point p of the zero velocity curve Z which can be followed through all three regions to reach $\theta = 0$ with $v = 0$ (the curve in figure 7). Using the reflection symmetries we get a periodic brake orbit, completing the proof of Theorem 2.

5.2. **Proof of Lemma 5.** The behavior of the stable and unstable manifolds on the isosceles triple collision manifold as the mass ratio m_3 varies has been studied using a combination of analytical and numerical methods by Simó ([23]). His results imply that the admissible masses, m_3, are those in the open interval $(0, 2.6620)$. In this section we will only prove that $m_3 = 1$ is admissible. The admissible masses form an open set so they will then include some open interval around 1. Using (θ, v) as coordinates on the collision manifold and using the energy equation with $r = 0$ in (45) gives

$$v' = \tfrac{1}{8} w^2 (1 + \sin^2\theta)^2$$
$$\theta' = \tfrac{1}{4} w (1 + \sin^2\theta)^2.$$

Any orbit segment with $w > 0$ can be parametrized by θ and we have

$$\frac{dv}{d\theta} = \tfrac{1}{2} w = \frac{2\sqrt{2W(\theta) - v^2\cos^2\theta}}{1 + \sin^2\theta} = \frac{2|\cos\theta|\sqrt{2V(\theta) - v^2}}{1 + \sin^2\theta}. \tag{48}$$

Let $m_3 = 1$. We need to follow the $w > 0$ branches γ, γ' of the unstable manifolds $W^u(L_-), W^u(L'_-)$. These are solutions of (48) beginning at L_-, L'_- whose coordinates are $(\theta, v) = (\pi - \theta^*, -v^*), (-\theta^*, -v^*)$ where $v^* = \sqrt{2V(\theta^*)} = \sqrt{6}$.

The branch γ' is short and easy to understand using crude estimates. Let $v(\theta)$ denote the corresponding solution of (48). We have $v(-\theta^*) = -v^* = -\sqrt{6}$ and we need to show that $v(0) < 0$. In the interval $[-\theta^*, 0]$ we have $V(\theta) \le V(0) = \frac{5}{\sqrt{2}}$. From (48) we get $\frac{dv}{d\theta} \le \sqrt{2V(0)}$ and

$$v(0) \le -v^* + \sqrt{2V(0)}\,\theta^*.$$

For $m_3 = 1$ we have $\theta^* = \sin^{-1}(\sqrt{2} - 1)$. Evaluating these constants numerically gives $v(0) \le -1.3 < 0$ as required.

The branch γ requires more work. First, it must be shown that γ remains in the $w > 0$ part of the energy manifold at least until it crosses the line $\theta = 0$. Thus is equivalent to $v(\theta)$ being defined for $\theta \in [\theta^* - \pi, 0]$. Then we also require that that $v_1 = v(-\frac{\pi}{2}) < 0$ and $v_2 = v(0) > 0$. Recall that solutions can only leave $\{w > 0\}$ in region R_{II}. If $v(\theta)$ is defined for $\theta \in [\theta^* - \pi, -\frac{\pi}{2}]$ and if $v(-\frac{\pi}{2}) < 0$ then in R_{II}, γ is trapped between two branches of stable manifolds and cannot reach the boundary curve $\{w = 0\}$ (see figure 8). Thus it suffices to show that $v(-\frac{\pi}{2}) < 0$ and $v(0) > 0$.

To show that $v(-\frac{\pi}{2}) < 0$. Recall that the initial value is $v(\theta^* - \pi) = -v^*$ where $v^* = \sqrt{6} \approx 2.44949$. First we show $v(-\frac{3\pi}{4}) \le -1.6$ then we show that the change of $v(\theta)$ on $[-\frac{3\pi}{4}, -\frac{\pi}{2}]$ is at most 1.56. For the first part, use the estimate $V(\theta) \le V(-\frac{3\pi}{4}) < 4$ for $\theta \in [\theta^* - \pi, -\frac{3\pi}{4}]$. From (48) we have

$$\frac{dv}{d\theta} \le \frac{2|\cos\theta|\sqrt{8 - v^2}}{1 + \sin^2\theta}.$$

If $v(-\frac{3\pi}{4}) > -1.6$ we would have

$$\int_{-v^*}^{-1.6} \frac{dv}{\sqrt{8 - v^2}} < \int_{\theta^* - \pi}^{-\frac{3\pi}{4}} \frac{2|\cos\theta|\,d\theta}{1 + \sin^2\theta}.$$

Both sides can be integrated exactly and the resulting formulas evaluated numerically to arrive at the contradiction $0.4459 < 0.4456$. For $\theta \in [-\frac{3\pi}{4}, -\frac{\pi}{2}]$ we use the crude approximation

$$\frac{dv}{d\theta} \le \frac{2\sqrt{2W(\theta)}}{1 + \sin^2\theta}.$$

Symbolic differentiation of the right-side shows that it is strictly decreasing on $[-\pi, -\frac{\pi}{2}]$ which implies that its integral can be easily bounded using upper and lower Riemann sums. In particular, the upper sum for $[-\frac{3\pi}{4}, -\frac{\pi}{2}]$ with 100 steps gives an upper bound of $1.559 < 1.56$ for the change of v over this interval.

Finally, we show that $v(0) > 0$. From (48) we have

$$\frac{dv}{d\theta} \ge \frac{2|\cos\theta|\sqrt{v^{*2} - v^2}}{1 + \sin^2\theta}.$$

This can be used to show that $v(-\frac{\pi}{2}) \ge -\frac{v*}{\sqrt{2}}$. If this were false we would have

$$\int_{-v^*}^{-\frac{v*}{\sqrt{2}}} \frac{dv}{\sqrt{v^{*2} - v^2}} > \int_{\pi - \theta^*}^{-\frac{\pi}{2}} \frac{2|\cos\theta|\,d\theta}{1 + \sin^2\theta}.$$

The left side is $\frac{\pi}{4}$ and using $\sin(\theta^*) = -\sin(\pi - \theta^*) = \sqrt{2} - 1$ one finds that the right side is also $\frac{\pi}{4}$, a contradiction. Therefore, $v(\frac{\pi}{2}) \ge -\frac{v*}{\sqrt{2}}$.

Since $v(\theta)$ is increasing we have $v(\theta) \geq v(-\frac{\pi}{2}) \geq -\frac{v*}{\sqrt{2}} = -\sqrt{3}$ for $\theta \in [-\frac{\pi}{2}, 0]$. If we had $v(0) \leq 0$ then we would have $v(t)^2 \leq 3$ for $\theta \in [-\frac{\pi}{2}, 0]$. Then we would have

$$\frac{dv}{d\theta} \geq \frac{2|\cos\theta|\sqrt{v^{*2} - v(t)^2}}{1 + \sin^2\theta} \geq \frac{2|\cos\theta|\sqrt{3}}{1 + \sin^2\theta}.$$

Integration shows that $v(\theta)$ increases by at least $\frac{\pi}{2}\sqrt{3}$ for $\theta \in [-\frac{\pi}{2}, 0]$ giving the estimate

$$v(0) \geq v(-\tfrac{\pi}{2}) + \tfrac{\pi}{2}\sqrt{3} \geq (\tfrac{\pi}{2} - 1)\sqrt{3} > 0$$

a contradiction to the assumption that $v(0) \leq 0$. Hence $v(0) > 0$ as required.

6. Acknowledgements.

The authors would like to acknowledge helpful discussions and correspondences with Alan Weinstein and Alain Chenciner. We would also like to thank Greg Laughlin for pointing us to the story of the Pythagorean 3-body problem.

References

1. A. Albouy and A. Chenciner, *Le probléme des n corps et les distances mutuelles*, Inventiones mathematicae 131 pp. 151-184 (1998).
2. G.D. Birkhoff *Dynamical Systems*, American Mathematical Society Colloquium Publications, vol. **9** (1927).
3. C. Burrau *Numerische Berechnung eines Spezialfalles des Dreikorperproblems*, *Astronomische Nachrichten*, Band 195. Nr. 4662, 6, (1913), 114-118. `http://adsabs.harvard.edu/abs/1913AN....195..113B` .
4. 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
5. A. Chenciner and R. Montgomery, *A remarkable periodic solution of the three-body problem in the case of equal masses*, Ann. of Math., **152** (2000) no. 3, 881–901.
6. R. Easton, *Parabolic orbits for the planar three-body problem*, JDE, **52** (1984) 116–134.
7. R. Easton and R. McGehee, *Homoclinic phenomena for orbits doubly asymptotic to an invariant three-sphere*, Indiana Univ. Math. J., **28** (1979) 211–240.
8. G.M. Ewing *Calculus of Variations with Applications*, Dover Publications, Inc, New York, (1985).
9. J-L. Lagrange, *Essai sur le Probléme des Trois Corps.* Prix de l'Académie Royale des Sciences de Paris, tome IX, *in* volume 6 of œuvres (page 292), (1772).
10. C. Marchal, *How the minimization of action avoids singularities*, Celestial Mech. Dynam. Astronom. **83**, 325-354 (2002).
11. J. Marsden, *Lecture on Mechanics*, London Mathematical Society Lecture Notes Series **174**, Cambridge University Press (1992).
12. R. McGehee, *Triple collision in the collinear three-body problem*, Inv. Math, **27** (1974) 191–227.
13. R. McGehee, *A stable manifold theorem for degenerate fixed points with applications to celestial mechanics*, JDE, **14** (1973) 70–88.
14. K. Tanikawa and S. Mikkola, *A trial symbolic dynamics of the planar three-body*, (2008), arXiv:0802.2465v1.
15. R. Moeckel, *Orbits near triple collision in the three-body problem*, Indiana Univ. Math J.,, **32,4** (1983) 221–240.
16. R. Montgomery, *Infinitely many syzygies*, Arch. Rat. Mech.., **164, 4** (2002) 311–340.
17. R. Montgomery, *The zero angular momentum three-body problem: all but one solution has syzygies*, Erg.Th.Dyn.Sys., **27,6** (2007) 1933–1946.
18. R. Montgomery, *The N-body problem, the braid group, and action-minimizing periodic orbits*, Nonlinearity, **11,2** (1998) 363–376.
19. C. Moore, *Braids in Classical Gravity, Physical Review Letters 70, pp. 3675–3679, (1993)*

20. C. Robinson, *Homoclinic orbits and oscillation for the planar three-body problem*, JDE, **52** (1984) 356–377.
21. Otto Raul Ruiz, *Existence of Brake-Orbits in Finsler Mechanical Systems* U.C. Berkeley thesis in Mathematics, 1975
22. H. Seifert, *Periodische Bewegungen Mechanischer Systeme*, Math. Z. 51 (1948), transl. by W. McCain at `http://count.ucsc.edu/~rmont/papers/list.html` under the year 2006
23. C. Simó, *Analysis of triple collision in the isosceles three-body problem*, in *Classical Mechanics and Dynamical Systems*, Marcel Dekkar (1981), 203–224.
24. C. Simó, J. LLibre, *Charcterization of transversal homothetic solutions in the n-body problem*, *Arch. Rat. Mech.*, 77 (1981), 189–198.
25. C. Simó, R.Martinez, *Qualitative study of the planar isosceles three-body problem*, *Cel. Mech.*, 41.1-4, (1987/88), 179–251.
26. C.L. Siegel and J. Moser, *Lectures on Celestial Mechanics*, Springer-Verlag,New York (1971).
27. V. Szehebely*Burrau's problem of three bodies Proc. Nat. Acad. Sci.*, 58, (1967), 60-65. `http://adsabs.harvard.edu/abs/1967PNAS...58...60S`
28. V. Szehebely and F. Peters, *A new periodic solution to the problem of three bodies.* , The Astronomical Journal, v. 72, no.9, (1967),1187-1190. `http://adsabs.harvard.edu/abs/1967AJ.....72.1187S`
29. K Tanikawa and S. Mikkola, *A trial symbolic dynamics of the planar three-body problem*, arXiv:0802.2465
30. A. Weinstein, Normal modes for nonlinear Hamiltonian systems, Inv. Math. 20 (1973), 4757.

School of Mathematics, University of Minnesota, Minneapolis MN 55455

Dept. of Mathematics, University of California, Santa Cruz, Santa Cruz CA

Laboratoire d'Analyse non linéaire et gémétrie, Université d'Avignon, Avignon (FR)
E-mail address: `rick@math.umn.edu`
E-mail address: `rmont@count.ucsc.edu`
E-mail address: `andrea.venturelli@univ-avignon.fr`