Arch. Rational Mech. Anal. 164 (2002) 311–340
Digital Object Identifier (DOI) 10.1007/s00205-002-0211-z

# *Infinitely Many Syzygies*

RICHARD MONTGOMERY

*Communicated by* P. J. HOLMES

**Abstract**

We show that any bounded zero-angular-momentum solution of the Newtonian three-body problem suffers infinitely many syzygies (collinearities) provided that it does not suffer a triple collision. Our motivation comes from the dream of building a symbolic dynamics for the three-body problem. The proof relies on the conformal geometry of the shape sphere.

## 1. Introduction

### *1.1. Infinitely many syzygies*

A solution to the Newtonian three-body problem suffers a *syzygy*, or eclipse, when the three bodies, considered to be point masses, become collinear. The solution is *bounded* if the distances between bodies remains bounded by a fixed constant for all time.

**Theorem 1.** *Every bounded solution of the Newtonian three-body problem with zero angular momentum and no triple collision suffers infinitely many syzygies.*

Three-body solutions with zero angular momentum are necessarily planar, so the theorem is really about the the planar three-body problem. Mark Levi conjectured the theorem during a conversation with Richard Montgomery in 1998.

Binary collisions are regarded as syzygies for the purposes of the theorem. We recall that when a solution suffers a binary collision it can be analytically continued through the collision by means of the Levi-Civita regularization process see LEVI-CIVITA [LeviCiv21]. The only obstruction to infinite time existence for a three-body solution is triple collision (SUNDMAN [Sun12]): as long as the solution suffers no triple collision, it can be continued analytically in time.

Theorem 1 is false for the planar three-body problem if we omit the zero-angular-momentum hypothesis. The Lagrange solutions illustrate this. In these solutions the three bodies form an equilateral triangle at every instant, and hence they never suffer syzygies. Bounded Lagrange solutions with non-zero angular momentum exist for all time, and all mass distributions.

Theorem 1 should be compared with the theorems of DIACU [Dia89, Dia92]. In the first paper, Diacu proves that the set of planar initial conditions which lead to syzygy is open within the set of all initial conditions. In the second paper, Diacu considers the set of planar solutions which have the property that, if they suffer one syzygy, then they suffer infinitely many syzygies. He shows that this set of solutions has full measure within the set of all bounded solutions which suffer at least one syzygy, provided that the masses $m_1, m_2, m_3$ of the bodies lie within a certain subset $S$ of the three-dimensional space of mass distributions. This subset $S$ has positive measure and contains mass distributions in which one mass is much greater than the other two.

The possibility remains that the set of solutions to which Theorem 1 applies is empty! Numerical evidence suggests otherwise. In the particular case when the masses are all equal we know that this set of solutions is nonempty, since it contains the periodic figure-eight solution [Che00]. Because the eight is (numerically) KAM stable, there will be an open set of near-equal masses for which this set of solutions is nonempty.

### *1.2. Motivation*

Syzygies come in three types, labelled 1, 2, and 3 according to which body is in the middle. See Fig. 1.

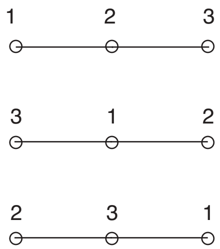


**Fig. 1.** The three syzygy types.

We can associate a syzygy sequence with each three-body motion, provided the solution is not collinear for all time, and provided it suffers no collisions. A syzygy sequence is a sequence in the letters 1, 2, and 3, listing the syzygies of a solution in order of occurrence. If the solution is periodic modulo rotations then its syzygy sequence is periodic. The free-homotopy type of curve which is periodic modulo rotation, whether it is a solution or not, is encoded by its (periodic) syzygy sequence. Does every periodic syzygy sequence arise as the syzygy sequence of some such solution? (Wu-Yi Hsiang asked this question in 1996. It helped lead to the rediscovery of the figure-eight solution of Chenciner & Montgomery [Che00].) More generally, is every infinite syzygy sequence realized by a solution? For various reasons, some given in Montgomery [Mon98], it is advantageous to restrict attention to zero-angular-momentum solutions. Is there a symbolic dynamics associated with the zero-angular-momentum three-body problem? The symbols would be 1, 2, 3 and perhaps the additional "halt symbols" 0 and $\infty$ to represent triple collision and escape to infinity. To build a symbolic dynamics, we mark occurences of syzygies in order. For this to work, we need some syzygies to occur. Theorem 1 asserts that syzygies occur infinitely often along bounded solutions with angular momentum zero and without triple collision.

## 2. Shape-space intuition and the evolution of spherical height

Newton's equations for the planar three-body problem are a system of six second-order differential equations. They reduce to a system of three second-order differential equations when we fix the values of the total linear and angular momenta and then divide out by the group of translations and rotations. These three equations describe evolution in *shape space*, this being the space of *oriented congruence classes of triangles*. Oriented congruence is distinguished from regular congruence in that two triangles related by a reflection are congruent but not oriented congruent. For example, shape space contains two distinct points $L_+$ and $L_-$ for the equilateral triangles of a fixed side length. One is "right-handed", and the other is "left-handed'. (We call these the *Lagrange points*.) See Fig. 2.

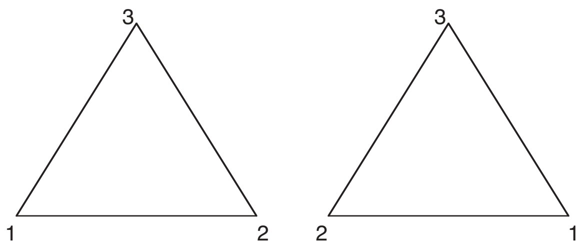


**Fig. 2.** The Lagrange configurations.

Shape space is homeomorphic to Euclidean three-space $\mathbb{R}^3$. See Fig. 3 below, and Section 4. The origin of this Euclidean space represents triple collision. Issuing from the origin are three rays parametrizing the binary collision configurations, one ray for each type of binary collision, mass 1 colliding with mass 2, mass 2 colliding with mass 3, or mass 3 colliding with mass 1. These rays are coplanar. The plane which they span is the syzygy plane, which parametrizes the space of collinear configurations.

A motion of the three bodies projects to the motion of a single point in shape space. If that motion is a zero-angular-momentum-solution to Newton's equation, then the motion of the point in shape space is governed by the system of three

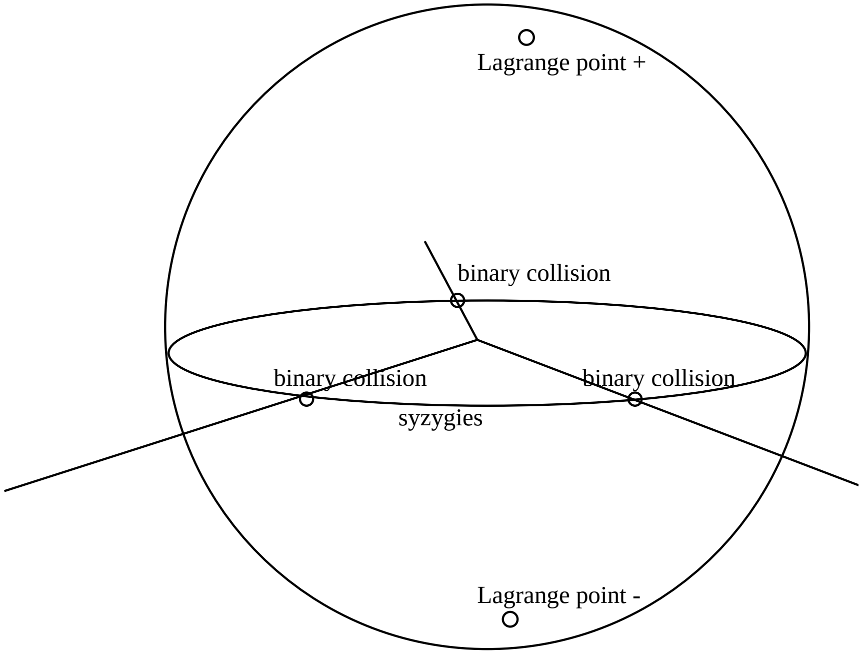


**Fig. 3.** Shape space.

second-order differential equation which comprise a system having the form of Newton's equations but now in shape space. These shape-space Newton's equations assert that each of the three binary collision rays exerts an attractive force on the point in shape space, and that the total force is the sum of these three forces. Since the three rays lie in the syzygy plane, this total force is always directed towards this plane. Based on this picture, Levi conjectured that the shape point is obliged to oscillate up and down forever, repeatedly crossing the syzygy plane, unless its distance from the plane tends to infinity.

We will turn Levi's intuition into a proof through the introduction of carefully chosen spherical coordinates $(R, \phi, \theta)$ on shape space. The radial coordinate $R$ measures distance from triple collision and is a measure of the overall size of a triangle. It is the square root of the moment of inertia as defined in (4.2.8). (See also (4.3.9), and (4.3.11).) Setting $R = 1$ defines a two-sphere in shape space which we identify with the *shape sphere*. The shape sphere is defined to be the space of oriented similarity classes of triangles. See Fig. 3 and Sections 4.3 and 4.4. The sphere is coordinatized by $\phi$ and $\theta$ (see (5.3)). The function

$$z = \sin(\phi)$$

is a measure of the height above the equator on the shape sphere and represents a normalized signed area of the triangle; see (5.4). The equator of the shape sphere is defined by $z = 0$ and represents the set of syzygy configurations. The "north and south poles", $z = 1$ and $z = -1$, represent the two Lagrange points, $L_+$

and $L_-$ which is to say the right-handed and left-handed equilateral triangles. The zero-angular-momentum Lagrange solutions, or *Lagrange homothety solutions* [Lag1772, Pol76] play a central role in the proof of Theorem 1. For such a solution $z$ is constant, $+1$ or $-1$ throughout. The solution consists of an equilateral-triangle configuration which shrinks by homothety to a point in finite time, ending in triple collision.

**Theorem 2.** *The normalized height coordinate* $z$ *(see* (5.4)*) on shape space minus triple collision, satisfies the following properties:*

(i) $-1 \leqq z \leqq 1$;
(ii) $z = \pm 1$ *if and only if the configuration is Lagrange* (*equilateral*)*; the functions* $1 - z$ *and* $-1 + z$ *measure the distance of a shape from the Lagrange shapes* $L_+, L_-$;
(iii) $z = 0$ *if and only if the configuration is a syzygy configuration.*

*Along any zero-angular-momentum solution to the three-body the function* $z$ *satisfies the differential equation*

$$\frac{d}{dt}(f\dot{z}) = -qz \qquad (*)$$

*away from triple collision. The functions* $f$ *and* $q$ *are non-negative and are given in* $(6.5b)$ *and* $(6.5c)$ *below. The function* $f$ *is a positive function on shape space. The function* $q$ *is a non-negative function on the tangent space to shape space whose zero locus coincides with the set of tangent vectors to the Lagrange homothety solutions.*

The proof of Theorem 2 is deferred to Sections 6 and 7.

## 3. Proof of Theorem 1.

We prove Theorem 1, assuming Theorem 2. We must prove that under the hypothesis of Theorem 1 the function $z$ of Theorem 2 has infinitely many zeros. Equivalently, we show that $z$ must have a zero on any infinite interval $[a, \infty)$.

Restrict attention to the case $z(t) > 0$. The argument for $z(t) < 0$ proceeds in an identical manner except that the signs of $z$ and its derivative $\dot{z}$ are to be reversed. We first show that if $z(t_1) > 0$ and $\dot{z}(t_1) < 0$, then at some later time $t_2 > t_1$ we must have $z(t_2) = 0$. Next we will show that if $z(t) > 0$, then eventually for some later time $t_* > t$ we must have $\dot{z}(t_*) < 0$. Together, these facts show that $z(t)$ has a zero some finite time later, and complete the proof.

So suppose that that $z(t_1) > 0$ and $\dot{z}(t_1) < 0$. Write $\dot{z} = \frac{1}{f}(f\dot{z})$ and integrate over the interval $t_1 \leqq s \leqq t$ to obtain

$$z(t) = z(t_1) + \int_{t_1}^{t} \frac{1}{f(s)}(f(s)\dot{z}(s))ds.$$

Set

$$\delta = -f(t_1)\dot{z}(t_1).$$

Since $f$ is positive, $\delta$ is positive. Differential equation (*) of Theorem 2, coupled with the non-negativity of $q$ says that that $f(s)\dot{z}(s)$ is monotone decreasing over any time interval on which $z$ is positive. That is, $f(s)\dot{z}(s) < f(t_1)\dot{z}(t_1) := -\delta < 0$ for $s > t_1$, as long as $z(s)$ is positive. The boundedness of our solution and the continuity of $f$ imply that $f$ is bounded along the solution. So there is a positive constant $K$ such that $0 < f(t) < K$ along our solution. Then $1/f > 1/K$ and $-1/f < -1/K$. Consequently $\dot{z} = (f\dot{z})/f < -\delta/K$ over our interval of positivity of $z$. Now suppose that $z(t)$ remains positive over the interval $t_1 \leqq s \leqq t_2$. It follows from our integral equation for $z(t)$ and the inequality immediately above that

$$z(t_2) < z(t_1) - (\delta/K)(t_2 - t_1).$$

This inequality together with $z(t) \leqq 1$ forces $z(t_2)$ to be negative as soon as $t_2 - t_1 > K/\delta$. Consequently $z$ must have a zero within the time interval $[t_1, t_1 + K/\delta]$.

It remains to show that there must be a time at which $\dot{z}$ is negative. This is equivalent to showing that it is impossible for a collision-free bounded zero-angular-momentum solution to simultaneously satisfy $z(t) > 0$ and $\dot{z} \geqq 0$ over an infinite time interval $a \leqq t < \infty$. We argue by contradiction. Suppose we have such a solution. Since $\dot{z} \geqq 0$ for all $t \geqq a$, the function $z$ is positive and monotone increasing over the whole infinite interval, and so tends to its supremum in infinite positive time. But $z$ is bounded by 1, so we must have $\dot{z} \to 0$. It follows that the limit of $f\dot{z}$ as $t \to \infty$ must be zero. (Again use the fact that $f$ is bounded along the solution.) We now show that the limit of $\lim_{t\to\infty} z(t) = 1$, which is to say, that the limiting shape is a Lagrange's equilateral triangle. For suppose this is not the case, then $z$ is everywhere positive and bounded away from the Lagrange shape, $z = 1$. Recall that the coefficient function $q$ of the differential equation (*) of Theorem 2 is non-negative and continuous, and is zero if and only if the shape is Lagrange and the initial conditions are those of a Lagrange homothety solution. It follows that if $\lim_t z(t) < 1$, then $q \geqq c$ everywhere along our solution, for some positive constant $c$. Now use the differential equation (*) of Theorem 2: $\frac{d}{dt}(f\dot{z}) = -qz$. Since $q \geqq c > 0$ and $z > z(a) > 0$, the right-hand side $-qz$ of this differential equation is strictly negative and bounded away from zero by the negative constant $-cz(a)$. This contradicts $\lim_{t\to\infty} f\dot{z} = 0$.

Now we know that $z \to 1$ monotonically as $t \to \infty$ while $f\dot{z}$ decreases monotonically to zero. The first fact says the configuration approaches the Lagrange equilateral shape. We will now show that there is a sequence of times $t_j$ tending to infinity for which the corresponding velocities approach those of the Lagrange homothety solution. Integrating the differential equation (*) of Theorem 2 from $t = a$ to $\infty$ and using $\lim_{t\to\infty} f(t)\dot{z}(t) = 0$ we obtain $\int_a^\infty q(s)z(s)ds = -f(a)\dot{z}(a)$. It follows that $\int_a^\infty q(s)ds$ is finite. This implies that the lim inf of $q$ as $t \to \infty$ is 0. Thus there are time intervals $[t_j, t_{j+1}]$, $t_j \to \infty$ over which $q(s)$ is as small as we please. (We have not excluded the possibility that $\limsup_{t\to\infty} q(t) > 0$.) During these intervals of small $q$ the solution is nearly tangent to the Lagrange homothety configuration, since this is the only place in phase space where $q$ is zero. In other words, the $\omega$-limit set of our solution curve contains points of phase space which are initial conditions for the Lagrange homothety solution.

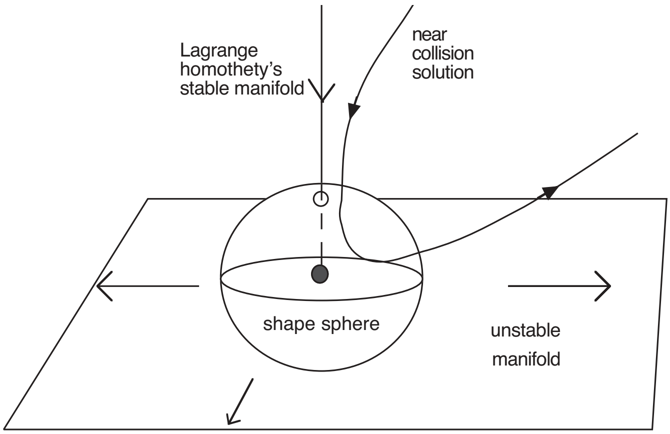


**Fig. 4.** A near-collision orbit.

It follows that our collision-free solution contains arcs which follow the Lagrange homothety solution arbitrarily closely, and hence come arbitrarily close to the Lagrange triple collision. We now use the results of MOECKEL [Moe83] on the linearization of the flow near Lagrange triple collision. Moeckel performs a McGehee-type blow-up to add the triple collision states as a boundary to phase space. The Lagrange triple-collision point becomes a hyperbolic rest point of saddle type for the resulting vector field, and the Lagrange homothety solution lies in its stable manifold. We have seen that our solution curve comes arbitrarily close to this saddle point. The solution cannot lie on the stable manifold of the saddle point, since if it did it would suffer a triple collision. It follows that the solution curve has near-collision hyperbolic shaped arcs in which it closely follows the stable manifold of the saddle point, coming very close to that point, then veers away to follow the unstable manifold in order to exit a small neighborhood of the point. Consequently the distance in phase space of the solution from the saddle point first decreases, then increases as it moves away along the unstable manifold. See Fig. 4.

We now show that the spherical distance $1 - z$ in configuration space from the Lagrange point must also increase. Near triple collision the unstable manifold of the Lagrange point is transverse to the fibers of the projection (configuration, velocity) $\mapsto$ (configuration). This transversality follows from the same transversality for the negative eigenspace of the linearized flow at the Lagrange point (MOECKEL [Moe83], pp. 228–229). Consequently, $1 - z$ must be increasing, hence we must have $\dot{z} < 0$, as desired. □

The following is a corollary to the proof of Theorem 1.

**Corollary 1.** *The normalized height function $z(t)$ of a zero-angular-momentum solution, bounded or not, has exactly one criticial point between any two consecutive zeros (syzygies) and this critical point is nondegenerate. Specifically, if $t_1 < t_2$ are consecutive zeros of $z(t)$ and if $t_c$ is the critical point in between these zeros, then $z(t)$ is strictly monotonic on the subintervals $t_1 < t < t_c$ and $t_c < t < t_2$.*

**Proof of Corollary 1.** Consider again the case $z > 0$. We saw in the proof of Theorem 1 that once $\dot{z} < 0$ then $z$ continues to decrease monotonically until it crosses zero. Thus it can have only one local maximum, on one side of which it is monotone increasing and the other side of which it is monotone decreasing. At the maximum $\dot{z} = 0$, so (*) of Theorem 2 becomes $f\ddot{z} = -qz$ there. Since $f$, $z$ and $q$ are all positive, we have $\ddot{z} < 0$ at this maximum point. (The possibility $q = 0$ is avoided as this implies that the solution is a Lagrange homothety solution.) When $z < 0$ the argument runs in the same way, resulting in $\ddot{z} > 0$ at the unique critical point of $z$ in an interval where $z$ is negative. □

## 4. Shape-space geometry

### *4.1. Shape space proper*

In this (long) section we set up the geometric tools needed to derive Theorem 2, and we explain the use of the variables $z, \phi, \theta$ there. The reader may wish to consult [Che00], or [Mon96, Mon98, Mon01] for more details.

We begin with a careful construction of shape space. We will identify the set $\mathbb{C}$ of complex numbers with the Euclidean plane in which the bodies move. The complex number $x_i$ specifies the location of the $i$th body, $i = 1, 2$ or 3. We write

$$Q = \mathbb{C} \times \mathbb{C} \times \mathbb{C} = \text{planar three-body configuration space}$$

and we write points of $Q$ as

$$x = (x_1, x_2, x_3) \in Q \quad x_i \in \mathbb{C}.$$

Write

$$SE(2) = \text{group of rigid motions of the plane}$$

for the group generated by translations and rotations of the Euclidean plane. This group is the group of isometries of the plane which preserve orientations and it acts on the $i$th plane according to $x_i \mapsto ax_i + b$ where $a = e^{i\theta} \in \mathbb{C}$ is a unit complex number representing rotation and $b \in \mathbb{C}$ represents the translation.

**Definition 1.** *Shape space* is the topological quotient space $Q/SE(2)$.

Continuous functions on shape space are identified with $SE(2)$-invariant continuous functions on $Q$. The three distances

$$r_{ij} := |x_i - x_j| \tag{4.1.1}$$

are such functions. So is the *signed area* of the triangular configuration $x \in Q$,

$$\Delta = \frac{1}{2}(x_2 - x_1) \wedge (x_3 - x_1), \tag{4.1.2}$$

where we write $z \wedge w = Im(z\bar{w}) = xv - yu$ where $z = x + iy$, $w = u + iv \in \mathbb{C}$.

If, when forming shape space, we had divided $Q$ instead by the full group $E(2)$ of isometries of the plane which includes reflection, then the syzygy plane would have formed a boundary and $\Delta$ would not have been a function on this shape space. By dividing by the smaller group $SE(2) \subset E(2)$ we "desingularize" this boundary, allowing smooth passage through syzygy. Any reflection of the Euclidean plane acts on shape space by reflection across the syzygy plane. It acts on the invariant functions (4.1.1), (4.1.2) by $r_{ij} \mapsto r_{ij}$, while $\Delta \mapsto -\Delta$.

### *4.2. Lagrangians*

Newton's equations for the planar three-body problem are the Euler-Lagrange equations for the Lagrangian

$$L = \tfrac{1}{2}K + U \tag{4.2.1}$$

where

$$U = \frac{m_1 m_2}{r_{12}} + \frac{m_1 m_2}{r_{13}} + \frac{m_2 m_3}{r_{23}} \tag{4.2.2}$$

is the negative of the potential energy, and

$$\begin{aligned} K &= m_1|\dot{x}_1|^2 + m_2|\dot{x}_2|^2 + m_3|\dot{x}_3|^3 \\ &= \langle \dot{x}, \dot{x} \rangle_m \end{aligned} \tag{4.2.3}$$

is twice the kinetic energy. Here $\dot{x}_i$ is the velocity of the $i$th body, and $m_i > 0$ is its mass. In the second line of (4.2.3) we have introduced the inner product

$$\langle v, w \rangle_m := \sum_{a=1}^{3} m_a v_a \cdot w_a; \qquad m = (m_1, m_2, m_3) \tag{4.2.4}$$

on the configuration space $Q$. Here $v_a \cdot w_a = \mathrm{Re}(v_a \bar{w}_a)$ denotes the standard inner product of the two vectors $v_a, w_a$ in $\mathbb{R}^2 = \mathbb{C}$, and the subscript $m$ indicates parametric dependence on the mass distribution. The inner product (4.2.4) is called the *kinetic energy inner product*.

We can decompose (4.2.3) according to

$$K = K_{\mathrm{shape}} + |J|^2/I + |P|^2/M, \tag{4.2.5}$$

where

$$P = \sum m_a \dot{x}_a = \text{ total linear momentum,} \tag{4.2.6}$$

$$J = \sum m_a x_a \wedge \dot{x}_a = \langle x, i\dot{x} \rangle_m = \text{ total angular momentum,} \tag{4.2.7}$$

$$M = m_1 + m_2 + m_3 = \text{ total mass,}$$

$$\begin{aligned} I &= \frac{1}{M} \sum m_i m_j r_{ij}^2 = \text{ total moment of inertia,} \\ &= \langle x, x \rangle_m \text{ when } \sum m_a x_a = 0. \end{aligned} \tag{4.2.8}$$

(The equality between the two lines of (4.2.8) is attributed to LAGRANGE [Lag1772]; see also [Pol76].) The $P$-term of $K$ measures translational kinetic energy. The $J$-term measures rotational kinetic energy. The remaining term, $K_{\mathrm{shape}}$, measures kinetic energy due to "internal" shape changes and corresponds to a Riemannian metric on shape space described in the next section.

Without loss of generality we may take $P = 0$ by going to a Galilean reference frame moving at constant velocity $P/M$. If we hypothesize, as we do for Theorem 1, that $J = 0$ also, then the Lagrangian (4.2.1) becomes

$$L_{\text{shape}} = \tfrac{1}{2} K_{\text{shape}} + U, \tag{4.2.9}$$

which is a Lagrangian on shape space. It is the Lagrangian whose Euler-Lagrange equations govern the motion of the projection to shape space of a zero-angular-momentum solution to Newton's equations.

**Remark.** Fixing the value of $J$ in the Lagrangian, pushing the result down to shape space, and then forming the Euler-Lagrange equations on shape space will only yield the correct reduced equations of motion when $J = 0$. When $J \neq 0$ a "magnetic" or Coriolis force term must be added to get the correct reduced equations. See MARSDEN & RATIU [Mar94] for example.

### *4.3. Geometry of shape space*

The group $SE(2)$ of rigid motions acts on a configuration $x \in Q$ sweeping out an orbit. A velocity vector $\dot{x} \in Q$ at $x$ is perpendicular to this orbit if and only if $P = 0$ and $J = 0$. The subspace of such perpendicular vectors is naturally identified with the tangent space to shape space at the shape $[x]$. (We write $x \mapsto [x]$ for the map $Q \to C := Q/SE(2)$ for the map sending a triangular configuration to its shape.) The restriction of the kinetic energy inner product to this perpendicular subspace defines a Riemannian metric on the shape space.

**Riemannian quotients.** The preceding construction is a special instance of the following general construction. Suppose that a Lie group $G$ acts on a Riemannian manifold $Q$ by isometries and that this action is free near the point $x$, meaning that $gx = x$ if and only if $g = Id$. Then the quotient space $Q/G$ is a manifold near the corresponding "shape" $[x] \in Q/G$ and its tangent space at $[x]$ is canonically identified with $T_x(Gx)^{\perp} \subset T_x Q$, the orthogonal complement to the tangent space to the group orbit $Gx \subset Q$. This identification yields a Riemannian metric on the quotient space by restricting the inner product on $T_x Q$ to $T_x(Gx)^{\perp}$. Because $G$ acts by isometries, this inner product does not depend on the representative point $x$ of the orbit $[x]$. We will refer to $Q/G$ with this metric as the *Riemannian reduction*, or the *Riemannian quotient* of $Q$ by $G$. This quotient metric satisfies the following properties.

(i) Every geodesic on $Q/G$ is obtained by taking a geodesic on $Q$ which is orthogonal to the group action and projecting it to $Q/G$.

(ii) The distance function on $Q/G$ associated with the Riemannian metric is the *orbital distance metric*: the distance between two points of $Q/G$ is the distance between the corresponding orbits in $Q$ relative to the distance function on $Q$.

We will realize the Riemannian reduction of the three-body configuration space $Q$ by the group $G = SE(2)$ of rigid motions in two steps. First we divide by translations. Then we divide by rotations. In between these two steps we will need to discuss the notion of the cone over a Riemannian manifold.

**Dividing by Translations.** We use the Jacobi vectors:

$$\xi_1 = x_2 - x_1, \quad \xi_2 = x_3 - \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}. \tag{4.3.1}$$

These vectors are invariant under translation of the $x_i$. See Fig. 5.

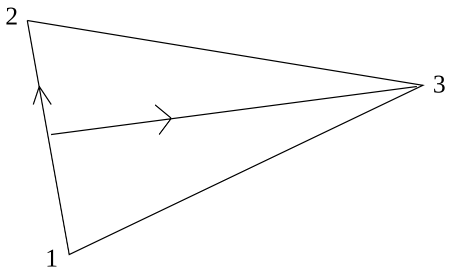


**Fig. 5.** The Jacobi vectors.

They diagonalize the kinetic energy, subject to the constraint that the total linear momentum is zero:

$$K = \mu_1 |\dot{\xi}_1|^2 + \mu_2 |\dot{\xi}_2|^2 \quad \text{provided } \sum m_i \dot{x}_i = 0.$$

Similarly

$$I = \mu_1 |\xi_1|^2 + \mu_2 |\xi_2|^2 \quad \text{provided } \sum m_i x_i = 0.$$

The coefficients $\mu_i$ are given by

$$\frac{1}{\mu_1} = \frac{1}{m_1} + \frac{1}{m_2}, \quad \frac{1}{\mu_2} = \frac{1}{m_3} + \frac{1}{m_1 + m_2}. \tag{4.3.2}$$

Set

$$z_1 = \sqrt{\mu_1}\xi_1, \quad z_2 = \sqrt{\mu_2} z_2. \tag{4.3.3}$$

Then

$$K = |\dot{z}_1|^2 + |\dot{z}_2|^2 \quad \text{provided } \sum m_i \dot{x}_i = 0, \tag{4.3.4}$$

$$I = |z_1|^2 + |z_2|^2 \quad \text{provided } \sum m_i x_i = 0. \tag{4.3.5}$$

The $z_i$ coordinatize the quotient vector space $Q$/translations. We call the $z_i$ *Jacobi coordinates*. They put the kinetic energy inner product (4.2.3) into standard form, that of the real part of the standard Hermitian inner product $((z_1, z_2), (w_1, w_2)) = Re(z_1 \bar{w}_1 + z_2 \bar{w}_2)$ on $\mathbb{C}^2$.

We summarize the change of variables $x \to \xi \to z$ with the mass-dependent *Jacobi map*

$$\mathcal{J}_m : Q \to \mathbb{C}^2; \quad \mathcal{J}_m(x_1, x_2, x_3) = (z_1, z_2). \tag{4.3.6}$$

The range, $\mathbb{C}^2$, of the Jacobi map realizes the quotient space of the configuration space by translations.

**Dividing by Rotations.** We proceed to the second step in the process of constructing shape space, which is to divide $Q$/translations $= \mathbb{C}^2$ by the action of rotations. Rotation by the angle $\theta$ acts on the Jacobi coordinates by complex scalar multiplication by $e^{i\theta}$: $(z_1, z_2) \mapsto (e^{i\theta} z_1, e^{i\theta} z_2)$. Thus shape space is isometric to the Riemannian quotient $\mathbb{C}^2/S^1$ where $\mathbb{C}^2$ is endowed with its standard Euclidean structure, the real part of the standard Hermitian form, and where the circle group $S^1$ acts by scalar multiplication on $\mathbb{C}^2$.

Following Hopf, we realize the quotient by this circle action by packaging the quadratic invariants for the circle action, namely $|z_1|^2$, $|z_2|^2$, $\mathrm{Re}(z_1\bar{z}_2)$, $\mathrm{Im}(z_1\bar{z}_2)$, into a single real vector

$$\mathbf{w} = (w_1, w_2, w_3) = \mathcal{H}(z_1, z_2) \tag{4.3.7a}$$

according to

$$w_1 = \tfrac{1}{2}(|z_1|^2 - |z_2|^2), \tag{4.3.7b}$$

$$w_2 + i w_3 = z_1 \bar{z}_2. \tag{4.3.7c}$$

The components $w_1$, $w_2$, $w_3$ of $\mathbf{w}$ form a global coordinate system for shape space. To summarize:

$$\mathbf{w}(x) := \mathcal{H}(\mathcal{J}_m(x)) \tag{4.3.7d}$$

is invariant under the action of the group $SE(2)$ of rigid motions on $Q$, and so induces a map from $Q/SE(2)$ to $\mathbb{R}^3$. This induced map is a homeomorphism of the quotient space onto $\mathbb{R}^3$.

Recall the signed area, $\Delta$ (equation (4.1.2)) and the moment of inertia $I$ (equations (4.2.8)). We compute that

$$w_3 = 4\sqrt{\frac{m_1 m_2 m_3}{m_1 + m_2 + m_3}}\,\Delta. \tag{4.3.8}$$

and that

$$\|\mathbf{w}\|^2 := w_1^2 + w_2^2 + w_3^2 = \tfrac{1}{4} I^2. \tag{4.3.9}$$

Fixing $I = 1$ defines a three-sphere $S^3$ in the space $Q$/translations $= \mathbb{C}^2$. The restriction of $\mathcal{H}$ to this three-sphere is the Hopf fibration, $S^3 \to S^2$, from the sphere of radius 1 in $\mathbb{C}^2$ onto the sphere of radius $1/2$ in $\mathbb{R}^3$. We call this two-sphere the *shape sphere*. Its points represent oriented similarity classes of triangles.

Introduce spherical coordinates $(\chi, \psi)$ according to

$$\frac{(w_1, w_2, w_3)}{\|\mathbf{w}\|} := (\cos(\psi)\cos(\chi), \sin(\psi)\cos(\chi), \sin(\chi)). \tag{4.3.10}$$

And set

$$R = \sqrt{I} \quad \text{so that} \quad R = \sqrt{2}\sqrt{\|w\|}. \tag{4.3.11}$$

**Lemma 1.** *The metric on shape space is*

$$ds^2_{\text{shape}} = dR^2 + \frac{R^2}{4} d^2 s_m \quad \textit{where} \quad d^2 s_m = d\chi^2 + \cos^2(\chi) d\psi^2 \qquad (4.3.12)$$

*so that twice the kinetic energy on shape space is*

$$K_{\text{shape}} = \dot{R}^2 + \frac{R^2}{4}(\dot{\chi}^2 + \cos^2(\chi)\dot{\psi}^2). \qquad (4.3.13)$$

*The variable $\chi$ is a normalized (signed) spherical distance from the syzygy plane in the shape sphere. It is zero if and only if we are at syzygy.*

The metric $d^2 s_m$ in Lemma 1 is the standard metric on the two-sphere of radius 1 in Euclidean space, so that the metric $d^2 s_m/4$ is the standard metric of radius 1/2 on the two-sphere. The metric $ds^2_{\text{shape}}$ on shape space (see (4.3.11),(4.3.12)) is that of the *cone over the two-sphere of radius* $1/2$, according to the following definition.

Let $X$ be a Riemannian manifold. Write the typical point in the product $[0, \infty) \times X$ of a ray with $X$ as $(R, x)$, where $R$ is a non-negative real number and $x$ is a point in $X$. The topological cone over $X$ is the quotient space $C(X) = ([0, \infty) \times X)/\sim$ where the identification "$\sim$" identifies all points having $R = 0$ to a single point, called the *cone point* and denoted 0. The cone has the structure of a smooth manifold away from the cone point.

**Definition (**The Cone over a Riemannian manifold**).** (See [Bur01] for the definition of a cone over a general metric space.) Let $d^2 s_X$ be the Riemannian metric on $X$. Put the Riemannian metric $ds^2 = dR^2 + R^2 d^2 s_X$ on the cone $C(X)$ over $X$. This is a smooth Riemannian metric away from the cone point. Its corresponding distance function extends continuously to the cone point, endowing the cone with the structure of a metric space.

The distance from the cone point is measured by $R$. The subset $\{R = 1\}$ is isometric to $(X, d^2 s_X)$.

*Example 1.* A Euclidean vector space $\mathbb{E}^{n+1}$ of dimension $n+1$ is isometric to the cone over the $n$-sphere $S^n(1) \subset \mathbb{E}^{n+1}$ of radius 1. The isometry $C(S^n(1)) \to \mathbb{E}^{n+1}$ sends $(R, \omega)$ to $R\omega \in \mathbb{E}^{n+1}$.

*Example 2.* Suppose that the Lie group $G$ acts on the Euclidean space $\mathbb{E}^{n+1}$ of E.1 by linear isometries. Then $G$ also acts on the sphere $S^n$ and so we can form the metric quotient $S^n/G$. Since the action of $G$ commutes with the action of dilations, the radial coordinates map $C(S^n(1)) \to \mathbb{E}^{n+1}$ of Example 1 is a $G$-equivariant map. It follows that $\mathbb{E}^{n+1}/G$ is isometric to the cone $C(S^n(1)/G)$ over $S^n(1)/G$. The coordinate $R$ on the cone is the Euclidean distance form the origin in $\mathbb{E}^{n+1}$. If $K$ is another Lie group which acts linearly on $\mathbb{E}^{n+1}$ and which commutes with the action of $G$, then this action descends to an action on the cone $C(S^n(1)/G) \cong \mathbb{E}^{n+1}/G$ as an action by isometries which preserves the coordinate $R$.

**Proof of Lemma 1.** Our situation here fits squarely within Example 2. Take $\mathbb{E}^4 = \mathbb{C}^2$, $G = S^1$, and $K = SU(2)$. We have seen that the quotient of $S^3(1)$ is the two-sphere and that the quotient map is realized by the Hopf fibration $S^3(1) \to S^2$, so that $\mathbb{C}^2/S^1 = C(S^2)$. It remains to figure out the metric on $S^2$. The Hopf fibration, viewed as a map $\mathbb{C}^2 \to \mathbb{R}^3$ is equivariant, where $SU(2)$ acts on $\mathbb{R}^3$ by the adjoint action, which is to say, via the $2:1$ cover $SU(2) \to SO(3)$. It follows that the metric on the image sphere $S^2$ of the Hopf fibration is homogeneous under $SO(3)$, and hence is a multiple of the standard metric on $S^2 = S^2(1)$. To see that the multiple gives the sphere a radius of $1/2$, note that points $(z_1, z_2)$ and $(-z_1, -z_2)$ are antipodal on $S^3(1)$ and hence every minimizing geodesic connnecting them has length $\pi$. However, $-1 \in S^1$ represents rotation by 180 degrees and hence $(z_1, z_2)$ and $(-z_1, -z_2)$ represent the same point in shape space. It follows that the projection to the shape space of a horizontal geodesic connecting these antipodal points is a full great circle whose circumference is $\pi$. But the circumference of a circle is $2\pi r$ where $r$ is the radius of the circle. Hence $r = 1/2$, accounting for the factor of $1/2$. □

We will need a formula expressing the distance $r_{ij}$ between bodies $i$ and $j$ in terms of the Hopf vector $\mathbf{w}$. Each type of binary collision defines a ray in $w$-space. Let $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ be the corresponding unit vectors, labelled so that $\mathbf{b}_1$ represents the collision $r_{23} = 0$ etc. These three vectors lie in the syzygy plane, $w_3 = 0$. In the formula below, and from now on, $ijk$ is any permutation of 123. The dot product and norm on the right-hand side of this equation (4.3.14) are the standard dot product and norm on $\mathbb{R}^3$.

**Lemma 2 (**Distance formula**).** *The distance between body $i$ and body $j$ is given by*

$$\begin{aligned} r_{ij}^2 &= \frac{m_i + m_j}{m_i m_j}(\|\mathbf{w}\| - \mathbf{w}\cdot\mathbf{b}_k) \\ &= \frac{m_i + m_j}{2m_i m_j} I(1 - \cos(\chi)\gamma_k(\psi)), \end{aligned} \tag{4.3.14}$$

*where $\mathbf{w}\cdot\mathbf{b}_k = \|\mathbf{w}\|\cos(\chi)\gamma_k(\psi)$ so that $\gamma_k(\psi)$ is the cosine of the angle between the projection of $\mathbf{w}$ onto the syzygy plane and the unit vector $\mathbf{b}_k$.*

**Proof of Lemma 2.** Let $x = (x_1, x_2, x_3) \in Q$ be a given configuration whose projection to shape space is $\mathbf{w}$. Move $x_i, x_j$ together along the line segment joining them in $\mathbb{C} = \mathbb{R}^2$ while keeping $x_k$ fixed, thus obtaining a straight line segment $x(t)$ in $Q$ and corresponding curve $\mathbf{w}(t)$ in shape space. We parametrize $x(t)$ so that the motion is linear in $t$, and choose the velocities so that the total linear and angular momenta of the motion is zero, and so that the curve ends with $i$ and $j$ colliding. The resulting line segment is a minimizing geodesic in $Q$ everywhere perpendicular to the $SE(2)$ orbits. It follows from properties (i) and (ii) of the paragraph "Riemannian quotients" above that $\mathbf{w}(t)$ is a minimizing geodesic in shape space which connects the given shape $\mathbf{w}$ to the binary collision ray. We caculate that the line segment $x(t)$ hits the $ij$ binary collision subspace, $\{x : x_i = x_j\}$ of $Q$ orthogonally, and hence minimizes the distance to this subspace. It follows that the curve $\mathbf{w}(t)$ realizes the distance $d_{ij}$ between the shape $\mathbf{w}$ and the binary collision ray. Back up on $Q$, we compute the kinetic energy length of $x(t)$ to be $\sqrt{\mu_{ij}}r_{ij}$ where $\mu_{ij} = m_i m_j/(m_i + m_j)$. It follows that

$$d_{ij} = \mu_{ij} r_{ij}. \tag{4.3.15a}$$

We now use the conical structure of the metric. Return to the cone over a general Riemannian manifold $X$. Let $c \subset X$ be a smooth curve having length $\theta_* < \pi/2$ and endpoints $c_0$, $c_1$. Then $C(c) \subset C(X)$ is the cone over a closed line segment of length $\theta_*$. Such a cone is isometric to the closed sector of the Euclidean plane defined by the inequalities $0 \leqq \theta \leqq \theta_*$ in polar coordinates. Let $p \in C(c_1)$ be the point on the bounding ray through $c_1$ which is a distance $R$ from the cone point. In other words, $p$ is the point represented as $(R, c_1)$. Then the distance $d$ from $p$ to the other bounding ray $C(c_0)$ is, from high-school geometry, $R\sin(\theta_*)$. Apply these considerations to the spherical projection of $c(t) = \mathbf{w}(t)/|\mathbf{w}(t)|$ of $\mathbf{w}(t)$ to obtain

$$d_{ij} = R\sin(\theta_{ij}), \tag{4.3.15b}$$

where $\theta_{ij}$ is the distance in $S^2(1/2)$ between $\mathbf{b}_k$ and the spherical projection of $\mathbf{w}$. Due to the radius of $1/2$, we have $\cos(2\theta_{ij}) = \mathbf{w}\cdot\mathbf{b}_k/|\mathbf{w}|$. Combine this with $|\mathbf{w}| = R^2/2$ (equation (4.3.9)) and the double-angle formula $\cos(2\theta) = 1 - 2\sin^2(\theta)$ to arrive at

$$|\mathbf{w}| - \mathbf{w}\cdot\mathbf{b}_k = R^2\sin^2(\theta_{ij}), \tag{4.3.16}$$

which combined with (4.3.15) yields the desired result, the first equation of (4.3.14). Writing the vector $\mathbf{w}$ in terms of the spherical coordinates $(R, \chi, \psi)$ and using $I = R^2$ yields the second equality of (4.3.14). □

## *4.4. Conformal geometry of the shape sphere*

We identified the shape sphere as the subspace $\{I = 1\}$ of the shape space. At a more fundamental level the shape sphere is the set of oriented similarity classes of triangles.

**Definition.** The *shape sphere* is the quotient space $Q^*/\tilde{G}$ where $Q^* = Q \setminus \{\text{triple collisions}\} \subset Q$ and where $\tilde{G} \supset SE(2)$ is the group of all orientation preserving similarity transformations of the Euclidean plane.

We perform the quotient of $Q^*$ by $\tilde{G}$ in two steps. First, we divide by translations so as to form

$$Q^*/\text{translations} \cong \mathbb{C}^2 \setminus \{0\}.$$

Next, we divide by the group of orientation-preserving similarities. The latter group acts on $\mathbb{C}^2 \setminus \{0\}$ as the group $\mathbb{C}^*$ of nonzero complex numbers acting by scalar multiplication. Thus

$$\text{shape sphere} \cong (\mathbb{C}^2 \setminus \{0\})/\mathbb{C}^* = \mathbb{CP}^1,$$

where $\mathbb{CP}^1$ is the complex projective line, the space of all complex lines in $\mathbb{C}^2$.

The metric we have put on the shape sphere depends on the choice $m$ of mass distributions. We wrote this metric as $d^2 s_m$ in Lemma 1 to indicate its mass dependence. A different mass distribution $m'$ will yield a different metric $d^2 s_{m'}$ on the same space, different despite the fact that the two metrics are isometric to each other. These two metrics are conformally related. We will need the precise conformal factor which relates them. This factor is given by Proposition 2 below.

To properly understand the situation, replace $\mathbb{C}^2$ by an abstract two-dimensional complex vector space $V$, where $V$ corresponds to our $Q$/translations. The shape sphere corresponds to the projectivization $\mathbb{P}V \cong \mathbb{CP}^1$ of $V$. A Hermitian inner product on $V$ determines a Riemannian metric on $\mathbb{P}V$. If $I : V \to \mathbb{R}$ denotes the square of the norm associated with this Hermitian inner product, we will write $d^2 s_I$ for the corresponding Riemannian metric on $\mathbb{P}V$. This Riemannian metric is defined by endowing the three-sphere $\{I = 1\} \subset V$ with the induced metric coming from the real part of the Hermitian inner product, and then identifying $\mathbb{P}V$ with the Riemannian quotient $\{I = 1\}/S^1$. If $(z_1, z_2)$ are $I$-orthonormal complex linear coordinates on $V$, so that $I = |z_1|^2 + |z_2|^2$, and if $z = z_1/z_2$ is the corresponding affine coordinate on $\mathbb{P}V$, then

$$ds_I = |dz|/(1 + |z|^2). \tag{4.4.1}$$

**Proposition 1.** *Let $I : V \to \mathbb{R}$ and $I' : V \to \mathbb{R}$ be the squared norms for two different Hermitian structures $\langle \cdot\,, \cdot \rangle$ and $\langle \cdot\,, \cdot \rangle'$ on the same complex two-dimensional vector space $V$. Let $\mathbb{CP}^1 = \mathbb{P}(V)$ be the projectivization of this vector space, and let $d^2 s_I$ and $d^2 s_{I'}$ be the two metrics on this complex projective line induced by our two Hermitian inner products. Let $L : V \to V$ be a linear operator intertwining the two norms: $\langle Lv, Lw \rangle = \langle v, w, \rangle'$. Then the two metrics are conformally related by*

$$ds_{I'} = |det(L)|(I/I')ds_I. \tag{4.4.2}$$

**Proof of Proposition 1.** Choose complex linear coordinates $z_1, z_2$ which are $I$-orthonormal and which diagonalize $I'$. Then

$$I = |z_1|^2 + |z_2|^2$$

while

$$I' = \lambda_1^2 |z_1|^2 + \lambda_2^2 |z_2|^2.$$

The linear map $L = \text{diag}(\lambda_1, \lambda_2)$ intertwines the two Hermitian structures. We know that

$$w_1 = \lambda_1 z_1; \quad w_2 = \lambda_2 z_2$$

are orthonormal for $I'$. Set $w = w_1/w_2$ and $z = z_1/z_2$. Then, according to (4.4.1) we have $ds_I = |dz|/(1 + |z|^2)$ and $ds_{I'} = |dw|/(1 + |w|^2)$. We also have

$$dw = \alpha dz \text{ where } \alpha = \lambda_1/\lambda_2.$$

Then

$$
\begin{aligned}
ds_{I'} &= \frac{|dw|}{|w|^2+1} \\
&= \frac{\alpha|dz|}{\alpha^2|z|^2+1} \\
&= \frac{|z|^2+1}{\alpha^2|z|^2+1}\frac{\alpha|dz|}{|z|^2+1} \\
&= \frac{|z_1|^2+|z_2|^2}{\alpha^2|z_1|^2+|z_2|^2}\frac{\alpha|dz|}{|z|^2+1} \\
&= \frac{|z_1|^2+|z_2|^2}{\lambda_1^2|z_1|^2+\lambda_2^2|z_2|^2}\frac{\lambda_1\lambda_2|dz|}{|z|^2+1} \\
&= \frac{I}{I'}\det(L)ds_I
\end{aligned}
$$

as desired. □

**Remark.** The intertwining operator $L$ is not unique, but it can easily be checked that if $F$ is another intertwining operator then $\det(F) = \det(L)$ so that the formula (4.4.2) is independent of the choice of intertwining operator, as it must be.

As a corollary to Proposition 1 we have

**Proposition 2.** *Let $m$ and $m'$ be two different mass distributions. Then the corresponding metrics $d^2s_m$ and $d^2s_{m'}$ on the shape sphere are conformally related according to the formula*

$$
\frac{m_1+m_2+m_3}{m_1m_2m_3}I_m^2d^2s_m = \frac{m_1'+m_2'+m_3'}{m_1'm_2'm_3'}I_{m'}^2d^2s_{m'}. \tag{4.4.3}
$$

**Proof of Proposition 2.** Proposition 1 tells us that $d^2s_{m'} = C(I_m^2/I_{m'}^2)d^2s_m$ and that the constant $C$ is given by $C = |\det(L)|^2$ where $L$ is an intertwining operator taking $I_m$ to $I_{m'}$. We find such an $L$ and compute its determinant.

Fix a triangular configuration $x = (x_1, x_2, x_3) \in Q$. It has two images, written $z$ and $w$, in $\mathbb{C}^2$ according to the Jacobi maps for the two different mass distributions $m$, and $m'$. Write $z = \mathcal{J}_m(x)$ and $w = \mathcal{J}_{m'}(x)$.

We look for a linear map $L : \mathbb{C}^2 \to \mathbb{C}^2$ such that $w = Lz$. Make the upper triangular anzatz $L(z_1, z_2) = (\alpha z_1, \beta z_1 + \gamma z_2)$. Using expressions (4.3.1)–(4.3.3) for the Jacobi map, the ansatz leads to the two linear equations $\alpha z_1 = w_1$ and $\beta z_1 + \gamma z_2 = w_2$, or

$$
\alpha\sqrt{\mu_1}(x_2-x_1) = \sqrt{\mu_1'}(x_2-x_1),
$$

and

$$
\beta\sqrt{\mu_1}(x_2-x_1)+\gamma\sqrt{\mu_2}(x_3-(m_1x_1+m_2x_2)/(m_1+m_2)) \\
= \sqrt{\mu_2'}(x_3-(m_1'x_1+m_2'x_2)/(m_1'+m_2')).
$$

The first equation has $\alpha = \sqrt{\mu_1'/\mu_1}$ for a solution. Expanding the second equation in $x_1, x_2, x_3$ and equating coefficients yields a system of three homogeneous equations, in the two unknowns $\beta$ and $\gamma$. The equation arising from the $x_3$ coefficient has $\gamma = \sqrt{\mu_2'/\mu_2}$ as a solution. Using this $\gamma$, the $x_1$ equation has $\beta = -\sqrt{\mu_2'/\mu_1}(m_1/(m_1+m_2) - m_1'/(m_1'+m_2'))$ for a solution, while the $x_2$ equation has $\beta = \sqrt{\mu_2'/\mu_1}(m_2/(m_1+m_2) - m_2'/(m_1'+m_2'))$ for a solution. These two $\beta$'s are checked to be equal, and so we get our invertible linear operator

$$L = \begin{pmatrix} \alpha & 0 \\ \beta & \gamma \end{pmatrix}.$$

We have $\det(L) = \alpha\gamma = \sqrt{\mu_1'\mu_2'/\mu_1\mu_2}$. Inserting in the formulae for the $\mu$ in terms of the masses leads to $\mu_1\mu_2 = m_1m_2m_3/(m_1+m_2+m_3) := c(m)$. Consequently $det(L) = \sqrt{c(m')/c(m)}$, the claimed result. □

The set of orientation-preserving conformal transformations of the two-sphere coincides with the group of orientation-preserving diffeomorphisms of the two-sphere which map circles to circles. We have just seen that the conformal structure of the shape sphere is independent of the mass distribution. It follows that the set of circles on the shape sphere can be specified in a manner independent of the choice of masses.

**Proposition 3.** *Every circle in the shape sphere is specified by a linear equation*

$$Ar_{12}^2 + Br_{23}^2 + Cr_{31}^2 + D\Delta = 0$$

*in the variables* $r_{12}^2, r_{23}^2, r_{31}^2, \Delta$ (*see* (4.1.1), (4.1.2)). *Such a circle passes through two triangles* (*points in the shape sphere*) *which are related by reflection if and only if* $D = 0$. *Such a circle passes through the two Lagrange configurations – the equilateral triangles* $L_+$, $L_-$ *– if and only if* $D = 0$ *and* $A + B + C = 0$.

**Minkowskian model of the conformal sphere.** To prove Proposition 3 we will use the Minkowsian model for the conformal structure of the two-sphere. Let $\mathbb{R}^{3,1}$ be a four-dimensional real vector space with a Minkowski inner product $\beta(\cdot,\cdot)$ of signature (3, 1). Choose linear coordinates so that $\beta(w, w) = -w_0^2 + w_1^2 + w_2^2 + w_3^2$. The subset $\{\beta(w, w) = 0, w_0 \geqq 0\}$ is called the positive light cone. Its real projectivization $\mathbb{P}C \subset \mathbb{P}(\mathbb{R}^{3,1})$ is the set of all forward-pointing light rays through the origin. It is topologically a two-sphere. The group of all invertible linear transformations of $\mathbb{R}^{3,1}$ which map $C$ to itself coincides with the group of time-oriented conformal Lorentz transformations – the group of linear transformations which preserve $\beta$ up to scale and which preserve the "sense of time" (sign of $w_0$). The identity component of this group will be denoted $CSO(3, 1)^+$. It acts by projective transformations on the sphere $\mathbb{P}C$. The circles on the sphere are the intersections of planes in $\mathbb{P}(\mathbb{R}^{3,1})$ (projectivized three-dimensional linear subspaces) with the sphere $\mathbb{P}C$. If $x_0, x_1, x_2, x_3$ are any linear coordinates on $\mathbb{R}^{3,1}$ then such a circle is described by a linear equation:

$$Ax_0 + Bx_1 + Cx_2 + Dx_3 = 0 \tag{4.4.4}$$

on the projectivized cone.

To see the isomorphism between the Minkowskian and complex Hermitian points of view regarding the conformal structure of the sphere, it helps to return to our abstract situation of a two-dimensional complex vector space $V$. The group $S^1$ of unit complex numbers acts on $V$ by complex scalar multiplication. We have seen that the quotient $V/S^1$ of $V$ by this group corresponds to shape space, and is homeomorphic to the cone over a two-sphere. Here is an alternative way to identify this same cone. Write $\mathcal{P}$ for the space of all real quadratic polynomials on $V$ which are invariant under the action of $S^1$. This is a four-dimensional vector space. If $z_1, z_2$ are linear coordinates on $V$, then a basis for $\mathcal{P}$ is formed by the polynomials $w_0 = \frac{1}{2}(|z_1|^2 + |z_2|^2)$, $w_1 = \frac{1}{2}(|z_1|^2 - |z_2|^2)$ $w_2 = \mathrm{Re}(z_1\bar{z}_2)$, and $w_3 = \mathrm{Im}(z_1\bar{z}_2)$. Write $\mathcal{P}^*$ for the real linear dual of $\mathcal{P}$. Define the map

$$\mathrm{ev} : V \to \mathcal{P}^*$$

("ev" for "evaluation") by

$$(\mathrm{ev}(v))(Q) = Q(v) \text{ for } v \in V,\ Q \in \mathcal{P}.$$

Since each function $Q \in \mathcal{P}$ is invariant under the circle action the map ev descends to the quotient by $S^1$, defining a map $V/S^1 \to \mathcal{P}^*$. This map from the quotient space is one-to-one and so its image is a faithful realization of the quotient space $V/S^1$. This image is the "postive light cone" $C = ev(V)$. In terms of our basis $\{w_0, w_1, w_2, w_3\}$ it is given by $-w_0^2 + w_1^2 + w_2^2 + w_3^2 = 0$ and $w_0 \geqq 0$. The cone is well defined independent of the choice of basis for $\mathcal{P}$, and itself defines on $\mathcal{P}^*$ a Lorentzian inner product ($\beta$) up to scale, and a time-orientation on $\mathcal{P}^*$, where $\mathcal{P}^*$ is our $\mathbb{R}^{3,1}$.

The map ev is $GL(V, \mathbb{C}) = GL(2, \mathbb{C})$ equivariant, where $GL(V, \mathbb{C})$ acts on the quadratic polynomials by pushforward: $g(Q)(v) = Q(g^{-1}v)$ and on $\mathcal{P}^*$ by the dual action. By construction, this action of $GL(V, \mathbb{C})$ on $\mathcal{P}^*$ preserves the quadratic cone $C$, and so is an action by conformal Lorentz transformations. Write $CSO(3, 1)^+$ for the group of conformal Lorentz transformations which are time-oriented and so send $C$ to itself. Thus the map ev defines a homomorphism

$$GL(V, \mathbb{C}) = GL(2, \mathbb{C}) \to CSO(3, 1)^+.$$

The kernel of this homomorphism is $\pm I$. Dimensional and connectivity considerations show that this homomorphism maps $GL(2, \mathbb{C})$ onto $CSO(3, 1)_+$.

To form the shape sphere we must divide by the action of dilations as well as rotations. These two actions commute. We have identified the quotient by rotations as the positive light cone $C \subset \mathcal{P}^*$. Dilation by $\lambda$ on $V$ acts by dilation by $\lambda^2$ on $\mathcal{P}^*$. Consequently, the shape sphere is identified with the real projectivized light cone, $\mathbb{P}C \subset \mathbb{P}(\mathbb{R}^{3,1})$. Then by (4.4.4) we know the circles on the shape sphere: they are given by any linear equation in $\mathcal{P}^*$ in the $w_i$.

**Proof of Proposition 3.** A basis for $\mathcal{P}$ is a linear coordinate system on $\mathcal{P}^*$. Instead of taking $w_0, w_1, w_2, w_3$ as basis we take the quadratic invariants

$$s_1 = r_{23}^2, s_2 = r_{31}^2, s_3 = r_{12}^2, \text{ and } \quad \Delta. \tag{4.4.6}$$

A circle on the shape sphere is then described, as in (4.4.4) above, by a linear equation $As_1 + Bs_2 + Cs_3 + D\Delta = 0$ in the $s_i$ and $\Delta$. This proves the first statement of the proposition. A triangle and its reflection have the same values for the invariants $s_i$ but their signed areas $\Delta$ are negatives of each other. So if a circle in the shape sphere passes through both a triangle and its reflection, then $As_1 + Bs_2 + Cs_3 + D\Delta = 0 = As_1 + Bs_2 + Cs_3 + -D\Delta$. Thus $D = 0$. This proves the second statement of the proposition. Finally, the two equilateral triangles $L_+$ and $L_-$ are related by reflection and satisfy $s_1 = s_2 = s_3$, which implies that $D = 0$ and that $A + B + C = 0$ for any circle passing through both. □

**Remark on Heron's formula.** The cone $C = \mathrm{ev}(V)$ can also be expressed in terms of the invariants $s_i, \Delta \in \mathcal{P}$. This expression is essentially Heron's formula for the area of a triangle:

$$\Delta^2 = p(p - r_{12})(p - r_{23})(p - r_{31}), \text{ with } p = \tfrac{1}{2}(r_{12} + r_{23} + r_{31}).$$

After a bit of algebra, Heron's formula becomes the quadratic relation

$$16\Delta^2 = 2s_1s_2 + 2s_3s_1 + 2s_2s_3 - (s_1^2 + s_2^2 + s_3^2),$$

which is the expression of the conformal Minkowskian inner product $\beta$ on $\mathcal{P}^*$ in the $\{s_i, \Delta\}$ basis. The time-orientation, or positive half $C$ of the light cone, is specified by adding the inequalities $s_i \geqq 0$, $i = 1, 2, 3$.

## 5. Good coordinates

We introduce the coordinates $(R, \phi, \theta)$ used for the proof of Theorem 2. Here $R$ is the usual square-root of the moment of inertia $I = I_m$ for our mass distribution $m = (m_1, m_2, m_3)$. The variables $\phi, \theta$ then coordinatize the shape sphere, but they are the metric spherical coordinates appropriate for the *equal mass distribution* $m' = (1, 1, 1)$ metric. Specifically, define the equal-mass Hopf coordinates

$$\mathbf{w} = \mathcal{H}(\mathcal{J}_{(1,1,1)}(x)) \tag{5.1}$$

(equations (4.3.7a–d)). Then

$$\|\mathbf{w}\| = \tfrac{1}{2} I_{(1,1,1)} = \tfrac{1}{2}\tfrac{1}{3}(r_{12}^2 + r_{23}^3 + r_{31}^2) \tag{5.2}$$

(equations (4.2.8) and (4.3.9)). The coordinates $\phi, \theta$ on shape space are defined by

$$\frac{1}{\|\mathbf{w}\|}\mathbf{w} = (\cos(\phi)\cos(\theta), \cos(\phi)\sin(\theta), \sin(\phi)). \tag{5.3}$$

The variable $z$ in Theorem 2 is

$$z = \sin(\phi). \tag{5.4a}$$

From (4.3.8) we know that

$$z = \frac{4}{\sqrt{3}} \frac{\Delta}{I_{(1,1,1)}} \tag{5.4b}$$

The variables $(\phi, \theta)$ are the earlier spherical coordinates $(\chi, \psi)$ of (4.3.10), (4.3.12) and (4.3.13), except that we have "artificially" set all masses equal to 1 in the formulae surrounding (4.3.10)–(4.3.13). According to Proposition 2,

$$d^2 s_{(1,1,1)} = d\phi^2 + \cos(\phi)^2 d\theta^2 \tag{5.5}$$

while

$$d^2 s_m = \lambda(\phi, \theta)^2 (d\phi^2 + \cos(\phi)^2 d\theta^2) \tag{5.6}$$

with

$$\lambda = \sqrt{\frac{3m_1 m_2 m_3}{M} \frac{I_{(1,1,1)}}{I}}. \tag{5.7}$$

Then, the kinetic energy on shape space for the mass distribution $m$ is

$$K_{\text{shape}} = \dot{R}^2 + \frac{R^2}{4} \lambda(\phi, \theta)^2 (\dot{\phi}^2 + \cos^2(\phi)\dot{\theta}^2). \tag{5.8}$$

For future use we set

$$K_{\text{sphere}} = \lambda(\phi, \theta)^2 (\dot{\phi}^2 + \cos^2(\phi)\dot{\theta}^2). \tag{5.9}$$

We have used the fact that $\lambda$, being a homogeneous $SE(2)$-invariant function of degree 0 can be expressed as a function of the spherical coordinates alone. The Lagrangian for the zero-angular-momentum three-body problem in our coordinates is

$$L_{\text{shape}} = \tfrac{1}{2} K_{\text{shape}} + U(R, \phi, \theta). \tag{5.10}$$

## 6. Proof of Theorem 2

It follows directly from the definition $z = \sin(\phi)$, (5.4a), that the normalized height variable satisfies properties (i) and (iii) of the theorem. When the masses are all equal, then the north and south poles of the sphere coincide with the Lagrange points, consequently $z$ satisfies property (ii) as well.

We proceed to derive the differential equation (*) of Theorem 2. We do this by computing the Euler-Lagrange equation for the evolution of the variable $\phi$ from the expression for the Lagrangian in the $(R, \phi, \theta)$ variables. The Euler-Lagrange equation for $\phi$ is $\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} = \frac{\partial L}{\partial \phi}$. Referring to (5.7)–(5.10) we see that

$$\frac{\partial L}{\partial \dot{\phi}} = \frac{R^2}{4} \lambda^2 \dot{\phi} \tag{6.1}$$

while

$$\frac{\partial L}{\partial \phi} = \frac{R^2}{4}\left[\lambda \frac{\partial \lambda}{\partial \phi}(\dot{\phi}^2 + \cos(\phi)^2\dot{\theta}^2) - \lambda^2 \cos(\phi)\sin(\phi)\dot{\theta}^2\right] + \frac{\partial U}{\partial \phi}, \tag{6.2}$$

so that the Euler-Lagrange equation for $\phi$ is

$$\frac{d}{dt}\left(\frac{R^2}{4}\lambda^2\dot{\phi}\right) = \frac{R^2}{4}\lambda\frac{\partial \lambda}{\partial \phi}(\dot{\phi}^2 + \cos(\phi)^2\dot{\theta}^2) - \frac{R^2}{4}\lambda^2\cos(\phi)\sin(\phi)\dot{\theta}^2 + \frac{\partial U}{\partial \phi}. \tag{6.3}$$

Now $\dot{z} = \cos(\phi)\dot{\phi}$, so that

$$\begin{aligned}\frac{d}{dt}\left(\frac{R^2}{4}\lambda^2\dot{z}\right) &= \cos(\phi)\frac{d}{dt}\left(\frac{R^2}{4}\lambda^2\dot{\phi}\right) + \left(\frac{R^2}{4}\lambda^2\dot{\phi}\right)\frac{d\cos(\phi)}{dt} \\ &= \cos(\phi)\frac{d}{dt}\left(\frac{R^2}{4}\lambda^2\dot{\phi}\right) - \frac{R^2}{4}\lambda^2\sin(\phi)\dot{\phi}^2.\end{aligned} \tag{6.4}$$

Substituting (6.3) into the first term of (6.4), using the expression (5.9) for $K_{\text{sphere}}$, the definition of $z$, and multiplying both sides of (6.4) by 4 we obtain:

$$\frac{d}{dt}(f\dot{z}) = -qz, \tag{6.5a}$$

where

$$f = R^2\lambda^2 = I\lambda^2, \tag{6.5b}$$

and

$$q = \left(1 - \frac{\cos(\phi)}{\sin(\phi)}\frac{1}{\lambda}\frac{\partial \lambda}{\partial \phi}\right)R^2 K_{\text{sphere}} - 4\frac{\cos(\phi)}{\sin(\phi)}\frac{\partial U}{\partial \phi}. \tag{6.5$c$}$$

This is the desired equation (*) of Theorem 2.

It is clear that $f > 0$ everywhere. It remains to establish the claimed positivity of $q$. This follows immediately from the following lemmas.

**Lemma 3.** *The following equality holds:*

$$\left(1 - \frac{\cos(\phi)}{\sin(\phi)}\frac{1}{\lambda}\frac{\partial \lambda}{\partial \phi}\right) = c/\lambda, \tag{6.6}$$

*where $c$ is the constant $\frac{1}{M}(m_1m_2 + m_3m_1 + m_2m_3)/\sqrt{\frac{3m_1m_2m_3}{M}}$, and where $\lambda$ is the conformal factor* (5.7).

**Lemma 4.** *The following inequality holds:*

$$-\frac{\cos(\phi)}{\sin(\phi)}\frac{\partial U}{\partial \phi} > 0 \tag{6.7}$$

*as long as $\phi \neq \pm\pi/2$, and $U$ is finite; i.e., everywhere except at the two Lagrange points and the three binary collision points of the shape sphere.*

This concludes the proof of Theorem 2, apart from the proofs of these two lemmas. □

**Remark on Lemma 4 at the equator.** It is not obvious at first glance that the functions appearing in lemmas 3 and 4 are smooth and well defined through the equator $\phi = 0$. Any function which is invariant under the full isometry group $E(2)$, including reflections, defines a function on shape space which is invariant under the reflection $\phi \to -\phi$ about the equator, and consequently is an even function of $\phi$ when $R, \theta$ are fixed. The functions $U$ and $\lambda$, being functions of the $r_{ij}$ alone, are such functions. Being even functions of $\phi$ their first partial derivative with respect to $\phi$ is an odd function of $\phi$, and these partial derivatives vanish at the equator $\phi = 0$. The function $\cot(\phi) = \cos(\phi)/\sin(\phi)$ is also an odd function of $\phi$, so both $\cot(\phi)\partial\lambda/\partial\phi$ and $\cot(\phi)\partial U/\partial\phi$ are even functions of $\phi$. A Taylor series analysis shows that at the equator $\phi = 0$ we have $\cot(\phi)\partial f/\partial\phi = \partial^2 f/\partial\phi^2$ for any smooth function $f = f(\phi, \theta, R)$ which is an even function of $\phi$. Thus Lemma 4 for $\phi = 0$ says that $-\partial^2 U/\partial\phi^2 > 0$.

## 7. Proof of the final lemmas

### *7.1. Proof of Lemma 3*

We introduce some simplifying notation for the purposes of this proof. Set

$$\hat{I} := I/I_{(1,1,1)}; \quad \hat{s}_k = r_{ij}^2/I_{(1,1,1)}; \quad p_k = m_i m_j/M \tag{7.1.1}$$

for $ijk$ a permutation of 123. Then

$$\hat{I} = \Sigma p_k \hat{s}_k, \tag{7.1.2}$$

while from (5.7)

$$\lambda = \sqrt{\frac{3m_1m_2m_3}{M}}\frac{1}{\hat{I}}. \tag{7.1.3}$$

Thus

$$-\frac{1}{\lambda}\frac{\partial\lambda}{\partial\phi} = +\frac{\partial}{\partial\phi}\log\hat{I}. \tag{7.1.4}$$

In this notation the claimed equality (6.6) is

$$1 + \frac{\cos(\phi)}{\sin(\phi)}\frac{\partial}{\partial\phi}\log\hat{I} = \sum p_k/\hat{I}. \tag{7.1.5}$$

We begin the verification of (7.1.5) by using formula (4.3.14):

$$r_{ij}^2 = I_{(1,1,1)}(1 - \gamma_k(\theta)\cos(\phi)).$$

(We have used $m_i = m_j = 1$ in (4.3.14), so that $(m_i + m_j)/2m_i m_j = 1$ there. Recall that our $\phi, \theta$ correspond to the $\chi, \psi$ of (4.3.10)–(4.3.14), but with all masses there set equal to 1.) It follows that

$$\hat{s}_k = 1 - \gamma_k(\theta)\cos(\phi). \tag{7.1.7}$$

From now on, write $\gamma_k$ for $\gamma_k(\theta)$. Then, from (7.1.7),

$$\frac{\partial \hat{s}_k}{\partial \phi} = \sin(\phi)\gamma_k$$

which, when substituted into the logarithmic derivative of (7.1.2), yields

$$\frac{\partial}{\partial \phi} \log \hat{I} = \sum p_k \sin(\phi)\gamma_k / \sum p_k \hat{s}_k.$$

Thus

$$\frac{\cos(\phi)}{\sin(\phi)} \frac{\partial}{\partial \phi} \log \hat{I} = \sum p_k \cos(\phi)\gamma_k / \sum p_k (1 - \gamma_k \cos(\phi)).$$

Finally

$$\begin{aligned} 1 &+ \frac{\cos(\phi)}{\sin(\phi)} \frac{\partial}{\partial \phi} \log \hat{I} \\ &= \frac{1}{\sum p_k (1 - \gamma_k \cos(\phi))} \left[ \sum p_k (1 - \gamma_k \cos(\phi)) + \sum p_k \gamma_k \cos(\phi) \right] \qquad (7.1.8) \\ &= \sum p_k / \sum p_k (1 - \gamma_k \cos(\phi)) \\ &= \sum p_k / \hat{I}. \end{aligned}$$

□

## 7.2. *Proof of Lemma 4*

Set

$$s_k = r_{ij}^2 \qquad (7.2.1)$$

for $ijk$ any permutation of 123. Then

$$U = m_1 m_2 / s_3^{1/2} + m_3 m_1 / s_2^{1/2} + m_2 m_3 / s_1^{1/2} \qquad (7.2.2)$$

while

$$I = (m_1 m_2 s_3 + m_3 m_1 s_2 + m_2 m_3 s_1)/M. \qquad (7.2.3)$$

Fixing the value of $R$ is the same as fixing that of $I$, since $I = R^2$. According to Proposition 3, fixing the value of $\theta$ is the same as imposing a linear constraint on the $s_i$:

$$As_1 + Bs_2 + Cs_3 = 0 \text{ with } A + B + C = 0. \qquad (7.2.4)$$

Thus, freezing $R$ and $\theta$ and varying $\phi$ is equivalent to imposing the two linear constraints (7.2.3) and (7.2.4) on the coordinates $s_k$ and varying the $s_k$ along the resulting line in the three-dimensional $s$-space. Since $1/s^{1/2}$ is convex for $s > 0$, and since $m_i > 0$, the potential function $U$ is a strictly convex function when restricted to the positive coordinate orthant $s_k > 0$ of $s_1 s_2 s_3$-space. A strictly convex function restricted to a line segment remains strictly convex. A strictly convex function has at most one local minimum, which is a global minimum when it exists. Our function $U$ has such a minimum when constraint (7.2.3) alone is imposed, and this minimum is the Lagrange point $s_1 = s_2 = s_3$, as is well known. The Lagrange point also satisfies constraint (7.2.4) according to Proposition 3, and so the Lagrange point remains the unique global minimum of $U$ when both the constraints (7.2.3) and (7.2.4) are imposed. Consequently, $U$ is strictly increasing as we move away from the Lagrange point on any of the circles $\theta = \text{const}$ of the shape sphere $R = \text{const}$. Now $\phi$ monotonically *decreases* as we move *away* from the positive Lagrange point toward the equator. This proves that

$$\frac{\partial U}{\partial \phi} < 0$$

for all $\phi$ with $0 < \phi < \pi/2$. Consequently, inequality of Lemma 3 holds in the upper hemisphere. It also holds in the lower hemisphere by reflectional symmetry.

Along the equator, we need a separate argument to establish the inequality. It is rather lengthy and is not used in the proof of Theorem 1, but we have decided to include it for the sake of completeness.

### *7.3. Positivity of $-\partial^2 U/\partial\phi^2$ along the equator*

According to the remark following the statement of Lemma 4, the positivity of $-\partial^2 U/\partial\phi^2$ along the equator $\phi = 0$ is equivalent to the claimed positivity $-\cot(\phi)\partial U/\partial\phi > 0$ at $\phi = 0$ (excepting the three binary collision points). To prove this positivity, we continue to use the notation of the proof of Lemma 3. There we used the notation $p_i = m_j m_k$ so that

$$U = M\left\{\frac{p_1}{s_1^{1/2}} + \frac{p_2}{s_2^{1/2}} + \frac{p_3}{s_3^{1/2}}\right\} \tag{7.3.1}$$

and

$$\frac{\partial U}{\partial \phi} = \frac{-M}{2}\left\{\frac{p_1}{s_1^{3/2}}\frac{\partial s_1}{\partial \phi} + \frac{p_2}{s_2^{3/2}}\frac{\partial s_2}{\partial \phi} + \frac{p_3}{s_3^{3/2}}\frac{\partial s_3}{\partial \phi}\right\}. \tag{7.3.2}$$

The functions $s_i = r_{jk}^2$ are even in $\phi$ so that $\partial s_1/\partial\phi = 0$ along the equator. It follows that, at the equator,

$$-\frac{\partial^2 U}{\partial \phi^2} = \frac{+M}{2}\left\{\frac{p_1}{s_1^{3/2}}\frac{\partial^2 s_1}{\partial \phi^2} + \frac{p_2}{s_2^{3/2}}\frac{\partial^2 s_2}{\partial \phi^2} + \frac{p_3}{s_3^{3/2}}\frac{\partial^2 s_3}{\partial \phi^2}\right\}. \tag{7.3.3}$$

We want to show that (7.3.3) is always positive along the equator, except at the binary collision points $s_i = 0$, the poles of $U$. We claim that it is enough to show that

**Lemma 5.** *Along the equator $\phi = 0$,*

$$s_i \leqq s_j < s_k \text{ implies } \frac{\partial^2 s_i}{\partial \phi^2} > 0 > \frac{\partial^2 s_k}{\partial \phi^2}. \tag{7.3.4}$$

We purposely wrote $s_j < s_k$ and not $s_j \leqq s_k$. For if $ij$ is the longest side, then, along the equator where the triangle is collinear, we have $r_{ij} = r_{ik} + r_{kj}$ from which it follows that $s_k := r_{ij}^2$ is strictly greater than both $s_i$ and $s_j$.

To see that Lemma 5 yields $-\partial^2 U/\partial\phi^2 > 0$ along the equator, recall our coordinates $R, \phi, \theta$. Now

$$R^2 = I = p_1 s_1 + p_2 s_2 + p_3 s_3 \tag{7.3.5}$$

so the partial derivative of $I$ with respect to $\phi$ is zero, as is its second derivative. Thus:

$$0 = p_1 \frac{\partial^2 s_1}{\partial\phi^2} + p_2 \frac{\partial^2 s_2}{\partial\phi^2} + p_3 \frac{\partial^2 s_3}{\partial\phi^2}. \tag{7.3.6}$$

Let us focus on the middle length which we suppose to be $r_{13} = \sqrt{s_2}$ without loss of generality. We will also suppose that the ordering is

$$s_1 \leqq s_2 < s_3. \tag{7.3.7}$$

Then, according to (7.3.4)

$$\frac{\partial^2 s_1}{\partial\phi^2} > 0 > \frac{\partial^2 s_3}{\partial\phi^2}. \tag{7.3.8}$$

Divide (7.3.6) by $s_2^{3/2}$ to obtain:

$$0 = \frac{p_1}{s_2^{3/2}} \frac{\partial^2 s_1}{\partial\phi^2} + \frac{p_2}{s_2^{3/2}} \frac{\partial^2 s_2}{\partial\phi^2} + \frac{p_3}{s_2^{3/2}} \frac{\partial^2 s_3}{\partial\phi^2}. \tag{7.3.9}$$

Now, from (7.3.7) we have $\frac{p_1}{s_1^{3/2}} \geqq \frac{p_1}{s_2^{3/2}}$ and, since $\frac{\partial^2 s_1}{\partial\phi^2}$ is positive, we have

$$\frac{p_1}{s_1^{3/2}} \frac{\partial^2 s_1}{\partial\phi^2} \geqq \frac{p_1}{s_2^{3/2}} \frac{\partial^2 s_1}{\partial\phi^2} > 0. \tag{7.3.10a}$$

Again from (7.3.7) we have $\frac{p_1}{s_2^{3/2}} > \frac{p_1}{s_3^{3/2}}$ and, since $\frac{\partial^2 s_3}{\partial\phi^2}$ is negative, we have

$$\frac{p_1}{s_2^{3/2}} \frac{\partial^2 s_3}{\partial\phi^2} < \frac{p_1}{s_3^{3/2}} \frac{\partial^2 s_3}{\partial\phi^2} < 0. \tag{7.3.10b}$$

Combining (7.3.9), with (7.3.10a,b) we see that indeed,

$$0 < \frac{p_1}{s_1^{3/2}} \frac{\partial^2 s_1}{\partial\phi^2} + \frac{p_2}{s_2^{3/2}} \frac{\partial^2 s_2}{\partial\phi^2} + \frac{p_3}{s_3^{3/2}} \frac{\partial^2 s_3}{\partial\phi^2}$$

which, according to (7.3.3), asserts that $-\partial^2 U/\partial\phi^2 > 0$ as desired.

It remains to establish Lemma 5.

**Proof of Lemma 5.** We continue to use the notation of Lemma 3:

$$\begin{aligned} s_k &= I_{(1,1,1)}\hat{s}_k \\ &= I\frac{I_{(1,1,1)}}{I}\hat{s}_k \\ &= I\frac{1}{\hat{I}}\hat{s}_k \end{aligned}$$

and

$$\hat{s}_k = 1 - \gamma_k \cos(\phi). \tag{7.3.11}$$

Recall that $I = R^2$ and $\gamma_k = \gamma_k(\theta)$ are constant as we vary $\phi$, and recall that $\hat{I}$ and $\hat{s}_k$ are even functions of $\phi$ so that $\partial\hat{I}/\partial\phi = \partial\hat{s}_k/\partial\phi = 0$ when $\phi = 0$. It follows that along the equator $\phi = 0$ we have:

$$\begin{aligned} \frac{\partial^2 s_k}{\partial\phi^2} &= I\left\{-\frac{1}{\hat{I}^2}\frac{\partial^2\hat{I}}{\partial\phi^2}\hat{s}_k + \frac{1}{\hat{I}}\frac{\partial^2\hat{s}_k}{\partial\phi^2}\right\} \\ &= \frac{I}{\hat{I}^2}\left[-\frac{\partial^2\hat{I}}{\partial\phi^2}\hat{s}_k + \hat{I}\frac{\partial^2\hat{s}_k}{\partial\phi^2}\right]. \end{aligned} \tag{7.3.12}$$

Since $I$ and $\hat{I}$ are both positive, it follows that the positivity or negativity of $\partial^2 s_k/\partial\phi^2$ agrees with that of $-(\partial^2\hat{I}/\partial\phi^2)\hat{s}_k + \hat{I}(\partial^2\hat{s}_k/\partial\phi^2)$. Now, from (7.3.11) we see that

$$\frac{\partial^2\hat{s}_k}{\partial\phi^2} = \gamma_k \tag{7.3.13}$$

along the equator. Here, and from now on, all partial derivatives are evaluated along the equator $\phi = 0$. Recalling (7.1.2), $\hat{I} = \Sigma p_k\hat{s}_k$, we see that

$$\frac{\partial^2\hat{I}}{\partial\phi^2} = \sum p_j\gamma_j.$$

Thus

$$-\frac{\partial^2\hat{I}}{\partial\phi^2}\hat{s}_k + \hat{I}\frac{\partial^2\hat{s}_k}{\partial\phi^2} = -\left(\sum p_j\gamma_j\right)\hat{s}_k + \left(\sum p_j\hat{s}_j\right)\hat{\gamma}_k.$$

The terms $p_k\hat{s}_k\gamma_k$ cancel, leaving us with

$$-\frac{\partial^2\hat{I}}{\partial\phi^2}\hat{s}_k + \hat{I}\frac{\partial^2\hat{s}_k}{\partial\phi^2} = p_i(\hat{s}_i\gamma_k - \gamma_i\hat{s}_k) + p_j(\hat{s}_j\gamma_k - \gamma_j\hat{s}_k) \tag{7.3.14}$$

for $ijk$ a permutation of 123.

Without loss of generality, we may assume that the ordering is

$$s_1 \leqq s_2 < s_3. \tag{7.3.15}$$

Then, according to (7.3.4), (7.3.12), and (7.3.14), $\partial^2 s_1/\partial\phi^2 > 0$ is equivalent to

$$p_2(\hat{s}_2\gamma_1 - \gamma_2\hat{s}_1) + p_3(\hat{s}_3\gamma_1 - \gamma_3\hat{s}_1) > 0 \tag{7.3.16a}$$

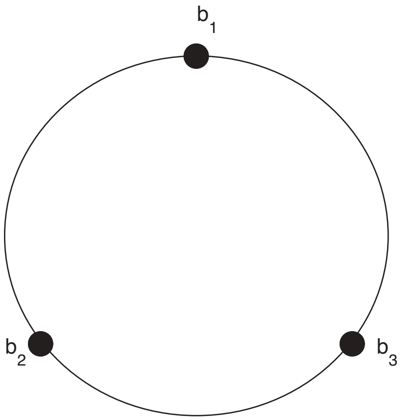


**Fig. 6.** The arrangement of binary collision configurations on the shape equator.

while $\partial^2 s_3/\partial\phi^2 < 0$ is equivalent to

$$p_1(\hat{s}_1\gamma_3 - \gamma_1\hat{s}_3) + p_2(\hat{s}_2\gamma_3 - \gamma_2\hat{s}_3) < 0. \tag{7.3.16b}$$

Now, at $\phi = 0$,

$$s_k = I_{(1,1,1)}\hat{s}_k = I_{(1,1,1)}(1 - \gamma_k),$$

so that

$$0 < \hat{s}_1 \leqq \hat{s}_2 < \hat{s}_3 \quad \text{and} \quad \gamma_1 \geqq \gamma_2 > \gamma_3. \tag{7.3.17}$$

Then we see in particular that

$$\gamma_1 - \gamma_2 \geqq 0 \text{ and } \gamma_1 - \gamma_3 > 0. \tag{7.3.18}$$

We also need to know

$$\gamma_3 < 0 < \gamma_1. \tag{7.3.19}$$

This follows directly from (7.3.17) and the fact that

$$\gamma_1 + \gamma_2 + \gamma_3 = 0. \tag{7.3.20}$$

To verify (7.3.20) we use the definition of our spherical coordinates in terms of the vector $\mathbf{w} = \mathcal{H}(\mathcal{J}_{(1,1,1)}(x))$. In these coordinates the locations of the binary collision points $\mathbf{b}_i$ on the shape sphere $I_{(1,1,1)} = 1$ are the vertices of an equilateral triangle inscribed in the equator. See Fig. 6.

In particular $\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3 = 0$. Taking dot products with the shape vector $\mathbf{w}$ yields (7.3.20), since $\cos(\phi)\gamma_i = \mathbf{w} \cdot \mathbf{b}_i$.

Combining (7.3.17) with (7.3.19), we see that

$$\hat{s}_3\gamma_1 > \hat{s}_2\gamma_1 \geqq \hat{s}_1\gamma_1 > 0 \tag{7.3.21}$$

while

$$\hat{s}_3\gamma_3 < \hat{s}_2\gamma_3 \leqq \hat{s}_1\gamma_3 < 0. \tag{7.3.22}$$

Combining (7.3.17), (7.3.18), and (7.3.21) we see that

$$\begin{aligned} p_2(\hat{s}_2\gamma_1 - \gamma_2\hat{s}_1) + p_3(\hat{s}_3\gamma_1 - \gamma_3\hat{s}_1) &> p_2(\hat{s}_1\gamma_1 - \gamma_2\hat{s}_1) + p_3(\hat{s}_1\gamma_1 - \gamma_3\hat{s}_1) \\ &= p_2\hat{s}_1(\gamma_1 - \gamma_2) + p_3\hat{s}_1(\gamma_1 - \gamma_3) \\ &> 0 \end{aligned}$$

which is the desired result, (7.3.16a), implying $\partial^2 s_1/\partial\phi^2 > 0$. A similar sequence of manipulations leads to (7.3.16b) and hence to $\partial^2 s_3/\partial\phi^2 > 0$ . (The sign of $\partial^2 s_2/\partial\phi^2$ can be either positive or negative.) □

## 8. Conclusion and open problems

We have shown (Theorem 1) that any bounded solution with zero angular momentum and no triple collision suffers infinitely many eclipses. But the set of such solutions may be empty!

**Open problem.** Show that for any mass distribution there exists a bounded solution with no triple collision and zero angular momentum.

The figure-eight solution ([Che00]) demonstrates that this set of solutions is non-empty when all three masses are equal.

The computations involved in the proof of Theorem 2 demonstrate that for certain applications our spherical three-body coordinates $R$, $\phi$, $\theta$ are optimal. They have also illustrated the power of the Albouy & Chenciner [Alb98] squared length coordinates $s_k = r_{ij}^2$ for shape space computations. We believe these coordinates will be of use in obtaining future results, including those with non-zero angular momentum.

*Acknowledgements.* I would like to thank Alain Albouy, Alain Chenciner, Mark Levi, Rick Moeckel and Jeff Xia for conversations crucial to the development of this paper. Alain Albouy insisted that using the squared lengths $r_{ij}^2$ would be the right way to establish Lemma 4. He also explained the Minkowskian geometry behind Proposition 3. Without these contributions the computations here would have been enormously more difficult. I would also like to thank the anonymous reviewer for a thorough and careful job that improved the paper. This work was supported in part by NSF grant (DMS 9704763).

## References

[Alb98] A. Albouy & A. Chenciner, Le probléme des n corps et les distances mutuelles, *Inventiones* **131** (1998), 151–184.

[Bur01] D. Burago, Y. Burago & S. Ivanov, *A course in metric geometry*, Graduate studies in Mathematics series **33**, AMS, 2001.

[Che00] A. Chenciner & R. Montgomery, A remarkable periodic solution of the three-body problem in the case of equal masses, *Annals of Mathematics*, **152** (2000), 881–901.

[Dia89] F. Diacu, On the planar syzygy solution of the 3-body problem, *Celestial Mechanics* **46** (1989), 119–128.

[Dia92] F. DIACU, A generic property of the bounded syzygy solutions, *Proc. of the AMS* **116** (1992), 809–812.

[Lag1772] J.-L. LAGRANGE, Essai sur le Probléme des Trois Corps. *Prix de l'Académie Royale des Sciences de Paris*, tome IX, p. 292 of volume 6 of œuvres, 1772.

[LeviCiv21] T. LEVI-CIVITA, Sur La Régularisation du Problème des Trois Corps, *Acta Math.* **42** (1921), 99–144.

[Mar94] MARSDEN, JERROLD E. & TUDOR S. RATIU, *Introduction to mechanics and symmetry*. Springer-Verlag, 1994.

[Moe83] R. MOECKEL, Orbits Near Triple Collision in the Three-Body Problem, *Indiana Univ. Math. J.*, **32** (1983), 221–240.

[Mon01] R. MONTGOMERY, A new solution to the three-body problem, *Notices of the AMS*, May (2001), 471–481.

[Mon98] R. MONTGOMERY, The $N$-body problem, the braid group, and action-minimizing orbits, *Nonlinearity* **11** (1998), 363–376.

[Mon96] R. MONTGOMERY, The geometric phase of the three-body problem, *Nonlinearity* **9** (1996), 1341–1360, esp. Section 1.2.

[Pol76] H. POLLARD, *Celestial mechanics*, Carus mathematical monographs no. 18, Mathematical Association of America, 1976.

[Sun12] K. SUNDMAN, Mémoire sur Le Problème des Trois Corps, *Acta Math.* **36** (1912), 105–179.

Mathematics Dept., UCSC
Santa Cruz
CA 95064, USA
e-mail: rmont@math.ucsc.edu

*(Accepted February 12, 2002)*
*Published online September 9, 2002 –* 