The Three-Body Problem and the Shape Sphere

Richard Montgomery

Abstract. The three-body problem defines dynamics on the space of triangles in the plane. The shape sphere is the moduli space of oriented similarity classes of planar triangles and lies inside *shape space*, a Euclidean three-space parameterizing oriented congruence classes of triangles. We derive and investigate the geometry and dynamics induced on these spaces by the three-body problem. We present two theorems concerning the three-body problem whose discovery was made through the shape space perspective.

1. INTRODUCTION. In 1667 Newton [**21**] posed the three-body problem. Central questions within the problem remain open today despite penetrating work over the centuries by many of our most celebrated mathematicians, including Euler, Lagrange, Laplace, Legendre, d'Alembert, Clairaut, Delanay, Poincare, Birkhoff, Seigel, Kolmogorov, Arnol'd, Moser, and Smale.

The problem, in its crudest form, asks us to solve the system of Ordinary Differential Equations [ODEs], Equations (1), governing the motion of three point masses moving in space under the sole influence of each other's mutual gravitational attractions. The masses form the vertices of a triangle, so we can think of the problem as one of moving triangles. According to the relativity principle of Galilieo, the laws of physics are invariant under isometries. Isometries are the congruences of Euclid. Two triangles are congruent if and only if the lengths of their three sides are equal. Is there a system of second order ODEs in the lengths of the three sides that describes the three-body problem?

All attempts to write down such a system of ODEs break down at collinear triangles. Instead, we will derive three alternative variables to the side lengths of a triangle and show that there is such a system of ODEs in these variables. Unlike the side lengths, the alternative variables are not invariant under all congruences but, rather, only under "oriented congruences." Two triangles are "oriented congruent" if a composition of a translation and rotation takes one to the other. Oriented congruence excludes reflections. We define shape space to be the space of oriented congruence classes of planar triangles. Shape space is homeomorphic to $\mathbb{R}^3$ and is parameterized by the vector (w_1, w_2, w_3) formed by these three alternative variables. We derive second order ODEs (Equation (54)) for these alternative variables that encode a special case of the three-body problem.

Shape space is homeomorphic to $\mathbb{R}^3$ but it is not isometric to $\mathbb{R}^3$. The shape space metric is not Euclidean. Nevertheless the shape space metric does enjoy spherical symmetry, which means that at the heart of shape space geometry is a two-dimensional sphere. We call this sphere the *shape sphere*. Its points represent *oriented similarity classes* of planar triangles (Figure 2). The purpose of this article is to describe the geometry of this sphere, how it relates to the three-body problem, and how this relation yields new insights into this age-old problem.

http://dx.doi.org/10.4169/amer.math.monthly.122.04.299
MSC: Primary 70F10, Secondary 70F15; 37N05

The shape sphere played a central role in our work on the three-body problem [**5**]. The shape sphere appeared previously in the three-body problem [**17**], and in statistics [**15**].

2. THREE-BODY DYNAMICS. Three point masses of magnitude m_1, m_2, m_3 move in space $\mathbb{R}^3$. Their positions as a function of time t are denoted by the position vectors $q_1(t), q_2(t), q_3(t) \in \mathbb{R}^3$. The three-body equations derived by Newton are

$$\begin{aligned} m_1\ddot{q}_1 &= F_{21} + F_{31}, \\ m_2\ddot{q}_2 &= F_{12} + F_{32}, \\ m_3\ddot{q}_3 &= F_{23} + F_{13}. \end{aligned} \tag{1}$$

We sometimes refer to the equations themselves as "the three-body problem." On the left-hand side of these equations the double dots mean two time derivatives: $\ddot{q} = \frac{d^2q}{dt^2}$. And below, a single dot over a variable will mean one time derivative. On the right-hand side of Equations (1)

$$F_{ij} = Gm_i m_j \frac{q_i - q_j}{r_{ij}^3}, \quad \text{where} \quad r_{ij} = |q_i - q_j| \tag{2}$$

is the force exerted by mass i on mass j. The constant G is Newton's gravitational constant and is physically needed to make dimensions match up. Being mathematicians, we set $G = 1$. The m_i are positive numbers. The Equations (1) are a system of second order equations in nine variables, the nine components of $\mathbf{q}(t) = (q_1(t), q_2(t), q_3(t)) \in \mathbb{R}^3 \oplus \mathbb{R}^3 \oplus \mathbb{R}^3$.

By design, the Equations (1) are invariant under the isometry group of space, that is, the group of transformations of $\mathbb{R}^3$ generated by

$$q \mapsto q + c\text{: translations,} \tag{3}$$

$$q \mapsto Rq \text{ : rotations,} \tag{4}$$

$$q \mapsto \bar{q} \text{ : reflections.} \tag{5}$$

In the first equation, $c \in \mathbb{R}^3$ is a translation vector. In the second equation, R is a rotation matrix: a three-by-three real matrix satisfying $RR^T = Id$ and $\det(R) = 1$. In the third equation, $q \mapsto \bar{q}$ is any reflection. (The full symmetry group of the Equations (1) is the Galilean group acting on space-time $\mathbb{R}^3 \oplus \mathbb{R}$, which strictly contains the group of isometries of space.)

Exercise 1. (A) Verify that the ODEs (1) are invariant under translation (3) as follows. Let $F : \mathbb{R}^3 \to \mathbb{R}^3$ be a translation: $F(q) = q + c$. Verify that if $\mathbf{q}(t) = (q_1(t), q_2(t), q_3(t))$ satisfies the Equations (1) then so does its translation: $F(\mathbf{q}(t)) := (F(q_1(t)), F(q_2(t)), F(q_3(t)))$.

(B) Formulate what it means for the Equations (1) to be invariant under rotations. Verify invariance.

(C) [Scaling]. Consider the space-time scaling transformation: $(q, t) \mapsto (\lambda q, \lambda^a t)$, $\lambda > 0$ that induces the action on curves: $q_i(t) \mapsto \lambda q_i(\lambda^{-a} t)$. Prove that the Equations (1) are invariant under this scaling transformation if and only if $a = 3/2$. Compare this scaling transformation with Kepler's third law.

(D) [Planar subproblem] Let $P \subset \mathbb{R}^3$ be a plane and $\mathbf{q}(t)$ a solution to the Equations (1) such that at some instant t_0 all three bodies lie in P while all three velocities are tangent to P. Show that all three bodies lie in P for all time t in the domain of the solution.

3. COMPLEX VARIABLES AND MASS METRIC. Exercise 1 (D) asserts under certain conditions the spatial three-body problem restricts to a plane, thus defining the "planar three-body problem." Choose xy axes for this plane P and then identify P with the complex number line $\mathbb{C}$ by sending a point $(x, y) \in P$ to the complex number $q = x + iy \in \mathbb{C}$. *The advantage of complex notation is that rotation corresponds to multiplication by a complex number of unit modulus.* In other words, we may replace the matrix formula (Equation (4)) for rotation by

$$q \mapsto uq, \quad u = \exp(i\theta)$$

where u is a unit complex number so that θ is real. The number θ is the radian measure of the amount of rotation. The set u of all unit complex numbers forms the circle group, denoted S^1.

We are now in the realm of Euclidean plane geometry. The locations $q_i \in \mathbb{C}$ of the three masses form the vertices of a Euclidean triangle.

Definition (Configuration space; located triangles). A planar triangle is represented by the vector

$$\mathbf{q} = (q_1, q_2, q_3) \in \mathbb{C}^3$$

whose components are the locations of the three vertices. The three-dimensional complex vector space $\mathbb{C}^3$ of all such three-vectors $\mathbf{q}$ is called the space of *located triangles*, or the *configuration space* for the three-body problem.

Definition (Mass metric). The mass metric on configuration space is the Hermitian inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle = m_1 \bar{v}_1 w_1 + m_2 \bar{v}_2 w_2 + m_3 \bar{v}_3 w_3. \tag{6}$$

Using the mass metric, we have that

$$K(\dot{\mathbf{q}}) = \frac{1}{2} \langle \dot{\mathbf{q}}, \dot{\mathbf{q}} \rangle := \frac{1}{2} \Sigma m_i |\dot{q}_i|^2 \tag{7}$$

is the usual kinetic energy of a motion. Here, $\dot{\mathbf{q}} = (\dot{q}_1, \dot{q}_2, \dot{q}_3) \in \mathbb{C}^3$ is the vector representing the velocities of the three masses. We also consider the gravitational potential energy

$$V(\mathbf{q}) = -\left\{ \frac{m_1 m_2}{r_{12}} + \frac{m_2 m_3}{r_{23}} + \frac{m_1 m_3}{r_{13}} \right\}. \tag{8}$$

Then

$$H(\mathbf{q}, \dot{\mathbf{q}}) = K(\dot{\mathbf{q}}) + V(\mathbf{q}), \tag{9}$$

is the energy of a motion $\mathbf{q}(t)$. In addition to the energy H, we will make use of several other important functions. The *moment of inertia*

$$I(\mathbf{q}) = \langle \mathbf{q}, \mathbf{q} \rangle = \Sigma m_i |q_i|^2 \tag{10}$$

measures the overall *size* of a located triangle. The *angular momentum*

$$J = Im(\langle \mathbf{q}, \dot{\mathbf{q}} \rangle) = m_1 q_1 \wedge \dot{q}_1 + m_2 q_2 \wedge \dot{q}_2 + m_3 q_2 \wedge \dot{q}_2, \tag{11}$$

measures the instantaneous spin of a triangle. The linear momentum

$$P = \langle \dot{\mathbf{q}}, \mathbf{1} \rangle = \Sigma m_i \dot{q}_i \in \mathbb{C} \tag{12}$$

measures the instantaneous rate of translation of the entire three-body system. The *center of mass* of the located triangle $\mathbf{q}$ is

$$q_{cm} = \langle \mathbf{q}, \mathbf{1} \rangle / \langle \mathbf{1}, \mathbf{1} \rangle = (m_1 q_1 + m_2 q_2 + m_3 q_3)/(m_1 + m_2 + m_3) \in \mathbb{C}. \tag{13}$$

In the formula for angular momentum, we used the notation

$$(x + iy) \wedge (u + iv) = det \begin{pmatrix} x & y \\ u & v \end{pmatrix} = xv - yu, \tag{14}$$

which is also equal to $Im(\bar{z}w)$ for $z = x + iy$, $w = u + iv \in \mathbb{C}$. This wedge operation $z, w \mapsto z \wedge w$ is the planar version of the cross product. If $\times$ denotes the usual cross product of vectors in $\mathbb{R}^3$, then $(x, y, 0) \times (u, v, 0) = (0, 0, z \wedge w)$ so that J is the third component of the usual angular momentum of physics. In the formulae for linear momentum and center of mass, we used the constant vector $\mathbf{1} = (1, 1, 1) \in \mathbb{C}^3$, which generates translations.

Definition (Phase space). The phase space of the planar three-body problem is $\mathbb{C}^3 \times \mathbb{C}^3$. Its points are written $(\mathbf{q}, \dot{\mathbf{q}})$. The first vector $\mathbf{q} \in \mathbb{C}^3$ represents the located triangle, that is to say, the positions of its three vertices. The second vector $\dot{\mathbf{q}}$ represents the velocities of these three vertices.

H and J are functions on phase space. I and q_{cm} are also functions on phase space but functions that are independent of $\dot{\mathbf{q}}$ and so can be thought of as functions on configuration space alone. The linear momentum is another function on phase space but now one that is independent of position $\mathbf{q}$.

Definition. A function $f : \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{R}$ or $F : \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}$ is *conserved* if its value is constant along any solution $\mathbf{q}(t)$ to the system of ODEs (1). (Different solutions typically yield different values for this constant.)

Proposition 1 (Conservation Laws). *The energy H, the angular momentum J, and the linear momentum P are* **conserved**.

The mass metric formalism yields a simple proof of this proposition. A complex vector space such as $\mathbb{C}^3$ becomes a real vector space when we allow only scalar multiplication by real scalars. The real part $\langle \cdot, \cdot \rangle_{\mathbb{R}}$ of the Hermitian mass inner product $\langle \cdot, \cdot \rangle$ of Equation (6) is a real inner product on $\mathbb{C}^3$. A real inner product induces a gradient operator that sends smooth real-valued functions $W : \mathbb{C}^3 \to \mathbb{R}$ to smooth real vector fields $\nabla W : \mathbb{C}^3 \to \mathbb{C}^3$ according to the rule

$$\frac{d}{d\epsilon} W(\mathbf{q} + \epsilon \mathbf{h})|_{\epsilon=0} = \langle \nabla W(\mathbf{q}), \mathbf{h} \rangle_{\mathbb{R}}. \tag{15}$$

In terms of real linear orthogonal (not necessarily orthonormal!) coordinates ξ^j, $j = 1, \ldots, 6$ for $\mathbb{C}^3$, the gradient ∇W is a variation of the usual coordinate formula from vector calculus. Namely, $(\nabla W)_j = \frac{1}{c_j} \frac{\partial W}{\partial \xi^j}$ where $c_j = \langle E_j, E_j \rangle$. Here, the linear coordinates ξ^j are related to an orthogonal basis E_j for $\mathbb{C}^3$ as per usual: $\mathbf{q} = \Sigma_{j=1}^6 \xi^j E_j$. We will take the ξ^j to come in pairs (x_j, y_j) as per $q_j = x_j + i y_j$ so that the c_j are then equal to the m_j in pairs, and we get the components of our gradient: $(\nabla V)_j = \frac{1}{m_j}\left(\frac{\partial V}{\partial x_j}, \frac{\partial V}{\partial y_j}\right) = \frac{1}{m_j}\left(\frac{\partial V}{\partial x_j} + i \frac{\partial V}{\partial y_j}\right)$.

Exercise 2. (A) Show that Newton's Equations (1) can be rewritten

$$\ddot{\mathbf{q}} = -\nabla V(\mathbf{q}) \quad \text{where } \nabla = \text{ gradient for the mass metric.} \tag{16}$$

(B) Use (A) to prove Proposition 1.

(C) Show that if $P = 0$ and $q_{cm}(0) = 0$, then $q_{cm}(t) = 0$ for all time t.

(D) The moment of inertia $I(t) = I(\mathbf{q}(t))$ evolves along a solution to Equations (1) according to the Lagrange–Jacobi equation:

$$\ddot{I} = 4H - 2V(\mathbf{q}).$$

4. THE TWO-BODY LIMIT. KEPLER'S PROBLEM. Set $m_3 = 0$ or $q_3 = \infty$ to kill the third equation of (1) and eliminate the variable q_3. The first two equations of (1) remain but now with $F_{31} = F_{32} = 0$. These remaining two equations form the "two-body problem." Set

$$\lambda = q_1 - q_2 \in \mathbb{C},$$

divide the first equation of (1) by m_1 and the second equation of (1) by m_2, and subtract it from the first to derive the single equation

$$\ddot{\lambda} = -c \frac{\lambda}{|\lambda|^3}, \tag{17}$$

with $c = m_1 + m_2$. This equation (for any $c > 0$) is often called "Kepler's problem," although Kepler did not write down any differential equations. Its solutions are the famous conics of Kepler's first law, parameterized according to Kepler's second law. The quantity

$$E = \frac{1}{2}|\dot{\lambda}|^2 - \frac{c}{|\lambda|}$$

is the associated conserved energy. A solution $\lambda(t)$ to Kepler's problem is bounded if and only if its energy E is negative. In this case, the curve described by λ is a circle or ellipse with one focus at $\lambda = 0$.

5. SOLUTIONS OF LAGRANGE AND JACOBI. Euler [**6**], followed quickly by Lagrange [**16**], wrote down the only solutions to the three-body problem for which we have explicit analytic expressions. Lagrange's solutions are depicted in Figure 1. To describe these solutions, note that we can rotate and scale a triangle by multiplying its vector $\mathbf{q}$ by a complex scalar

$$\mathbf{q} \mapsto \lambda\mathbf{q}, \lambda \in \mathbb{C}^* := \mathbb{C} \setminus 0 = \text{nonzero complex numbers.}$$

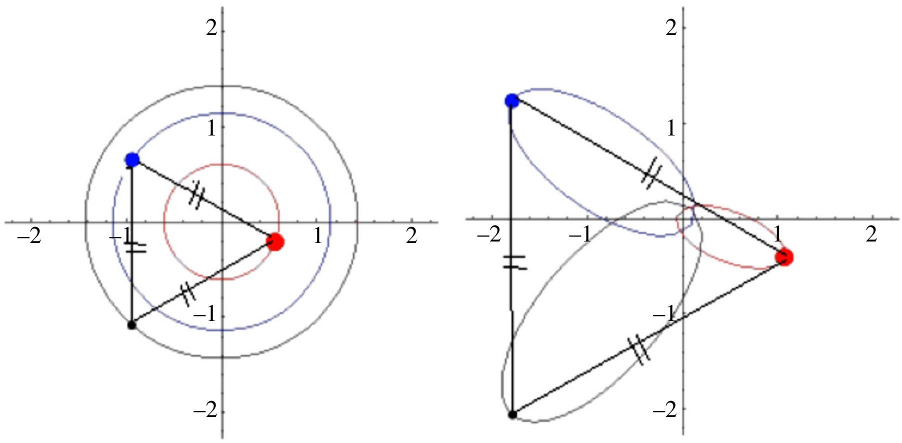

Figure 1. Lagrange solutions: The three bodies form an equilateral triangle at each instant. In the first figure, the triangle rotates about its center of mass so each individual orbit is a circle. In the second, the bodies travel on homothetic ellipses.

The magnitude $|\lambda|$ is the amount by which the triangle is scaled. The argument $\theta = Arg(\lambda)$ is the amount by which the triangle is rotated. Now suppose that a solution evolves solely by rotation and scaling

$$\mathbf{q}(t) = \lambda(t)\mathbf{q}_0, \quad \mathbf{q}_0 \neq 0, \quad \lambda(t) \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}. \tag{18}$$

Exercise 3. Show that the expression (18) solves Newton's Equations (1) if and only if the center of mass of $\mathbf{q}_0$ is zero and $\mathbf{q_0}$ and $\lambda(t)$ solve the two equations

$$\ddot{\lambda} = -c\frac{\lambda}{|\lambda|^3} \quad \text{and} \tag{19}$$

$$\frac{c}{2}\nabla I(\mathbf{q}_0) = \nabla V(\mathbf{q}_0) \tag{20}$$

where $c = -V(\mathbf{q}_0)/I(\mathbf{q}_0)$.

Hint: Use form (16) of Newton's equation, the equivariance identity $\nabla V(\lambda\mathbf{q}) = \frac{\lambda}{|\lambda|^3}\nabla V(\mathbf{q})$, and Euler's identity for homogeneous functions.

Equation (19) is the "Kepler problem" of the previous section. The second equation (20) is a Lagrange-multiplier-type equation asserting that $\mathbf{q}_0$ is a critical point of the function U, constrained to the sphere $I = I(\mathbf{q}_0)$. Modulo rotations and translations, there are exactly five such critical points, corresponding to the three collinear solutions found by Euler and the two equilateral solutions of Lagrange. They are represented by five points on the shape sphere (see Figure 2).

6. SHAPE SPACE. MAIN THEOREM. We seek reduced equations: ODEs encoding the three-body problem as dynamics on the space of congruence classes of triangles. Elementary geometry asserts that this space is three-dimensional with the triangle's side lengths serving as coordinates. So we expect ODEs in these lengths. However, the collinear triangles form the boundary of the space of congruence classes of triangles, and as a result, any ODE in side lengths alone will exhibit a singularity along the space of collinear triangles. But Newton's Equations (1) exhibit no such problem. The earth, moon, and sun line up without creating singularities in their dynamics. In order to achieve what we seek, we must strengthen the notion of congruence by excluding reflections as being allowed congruences.

Definition. The group G of rigid motions of the plane is the group of orientation-preserving isometries of the plane.

Every element of G is a composition of a rotation and a translation.

Definition. Two planar triangles (possibly degenerate) are "oriented congruent" if there is a rigid motion taking one triangle to the other.

Definition (Shape Space). Shape space is the space of oriented congruence classes of triangles, endowed with the quotient metric.

Replacing the equivalence relation of congruence by oriented congruence removes the boundary from the resulting space of equivalence classes (see Theorem 1 below).

Some words are in order regarding the meaning of "quotient metric" in the definition of shape space. The mass metric gives our space $\mathbb{C}^3$ of located triangles a norm $\|\cdot\|$ under which the distance between located triangles $\mathbf{q}, \mathbf{q}'$ is $\|\mathbf{q}-\mathbf{q}'\|$. G acts on $\mathbb{C}^3$ by isometries relative to this distance. Denote the result of applying $g \in G$ to $\mathbf{q} \in \mathbb{C}^3$ by $g\mathbf{q}$. We define the shape space metric d by

$$d([\mathbf{q}], [\mathbf{q}']) = \inf_{g_1, g_2 \in G} \|g_1\mathbf{q} - g_2\mathbf{q}'\|. \tag{21}$$

Here, $[\mathbf{q}]$ and $[\mathbf{q}']$ denote the "shapes," or oriented congruence classes, of the located triangles $\mathbf{q}$ and $\mathbf{q}' \in \mathbb{C}^3$.

Theorem 1 (Iwai [13]; references therein). *(See Figure 2.) Shape space is homeomorphic to $\mathbb{R}^3$. The quotient map from the space of located triangles to shape space is realized by a map $\pi : \mathbb{C}^3 \to \mathbb{R}^3$ that is the composition of a complex linear projection $\pi_{tr}\mathbb{C}^3 \to \mathbb{C}^2$ (Equations (24)) and a real quadratic homogeneous map $\pi^{rot} : \mathbb{C}^2 \to \mathbb{R}^3$ (Equation (31)). The map π enjoys the following properties.*

(A) *Two triangles $\mathbf{q}, \mathbf{q}' \in \mathbb{C}^3$ are oriented congruent if and only if $\pi(\mathbf{q}) = \pi(\mathbf{q}')$.*

(B) *π is onto.*

(C) *π projects the triple collision locus onto the origin.*

(D) *π projects the locus of collinear triangles onto the plane $w_3 = 0$ where (w_1, w_2, w_3) are standard linear coordinates on $\mathbb{R}^3$. Moreover, w_3 is the signed area of the corresponding triangle, up to a mass-dependent constant.*

(E) *Let $\sigma : \mathbb{R}^3 \to \mathbb{R}^3$ be the reflection across the collinear plane: $\sigma(w_1, w_2, w_3) = (w_1, w_2, -w_3)$. Then the two triangles $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{C}^3$ are congruent if and only if either $\pi(\mathbf{q}) = \pi(\mathbf{q}')$. or $\pi(\mathbf{q}) = \sigma(\pi(\mathbf{q}'))$.*

(F) $w_1^2 + w_2^2 + w_3^2 = (\frac{1}{2}I)^2$ *where* $I = \langle \mathbf{q}, \mathbf{q} \rangle$ *(Equation (10).*

(G) $R = \sqrt{I}$ *is the shape space distance to triple collision.*

Remark. D and E of the theorem say that the space of congruence classes of triangles can be identified with the closed half space $w_3 \geq 0$ of $\mathbb{R}^3$ and that the plane $w_3 = 0$ of collinear triangles forms its boundary, as claimed at the beginning of this section. Iwai [**13**] was, to my knowledge, the first to explicitly state something along the lines of this theorem. It has been independently discovered since then, and I would wager was known before Iwai also.

The metric and the shape sphere. Although shape space is homeomorphic to $\mathbb{R}^3$, it is not isometric to $\mathbb{R}^3$: Shape space geometry is not Euclidean. However, the geometry does have spherical symmetry. Each sphere centered at triple collision is isometric to the standard sphere, up to a scale factor. We identify these spheres with the *shape sphere*.

Add scalings in to the group G of rigid motions and we get the group of orientation-preserving similarities whose elements are compositions of rotations, translations, and scalings.

Definition. Two planar triangles are "oriented similar" if there is an orientation-preserving similarity taking one to the other.

Definition. The *shape sphere* is the resulting quotient space of the space of located triangles $\mathbb{C}^3 \setminus \mathbb{C}\mathbf{1}$ after the triple collisions $\mathbb{C}\mathbf{1}$ have been deleted.

In other words, the shape sphere is the space of oriented similarity classes of planar triangles where we do not allow all three vertices of the triangle to coincide. Now,

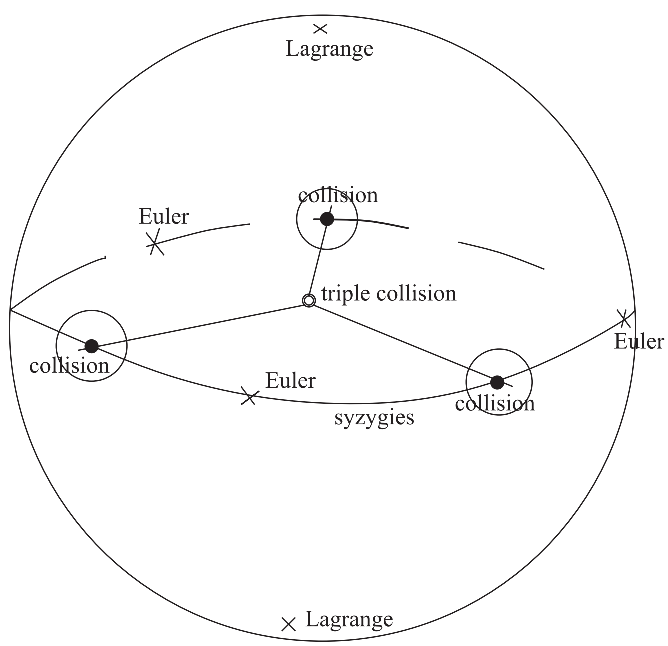

Figure 2. The shape sphere, centered on triple collision.

$\pi(\lambda\mathbf{q}) = \lambda^2\pi(\mathbf{q})$ for λ real. It follows that the shape sphere can be realized as the space of rays through the origin in $\mathbb{R}^3$. This space of rays can in turn be identified with the unit sphere $\|\mathbf{w}\| = 1$ within shape space. Various special types of triangles, including the five families of solutions of Euler and Lagrange, are encoded on this sphere as indicated in Figure 2.

7. FORMING THE QUOTIENT. PROVING THEOREM 1. A vector in $\mathbb{C}^3$ represents a located triangle with its three components representing the triangle's three vertices. Translation of a triangle $\mathbf{q} = (q_1, q_2, q_3) \in \mathbb{C}^3$ by $c \in \mathbb{C}$ sends $\mathbf{q}$ to the located triangle $\mathbf{q} + c\mathbf{1}$, where $\mathbf{1} = (1, 1, 1)$. Rotation by θ radians about the plane's origin sends $\mathbf{q}$ to $e^{i\theta}\mathbf{q} = (e^{i\theta}q_1, e^{i\theta}q_2, e^{i\theta}q_3)$. Scaling the plane by a positive factor ρ corresponds to multiplication by the real number ρ and so sends the triangle $\mathbf{q}$ to $\rho\mathbf{q} = (\rho q_1, \rho q_2, \rho q_3)$. See Figure 3.

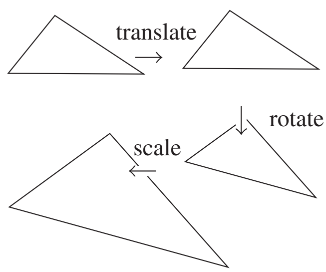

Figure 3. Translating, rotating, and scaling a triangle.

Shape space is the quotient of $\mathbb{C}^3$ by the action of the group G generated by translation and rotation. We form this quotient in two steps, translation then rotation.

Dividing by translations. We divide by translations by using the isomorphism

$$\mathbb{C}^3/\mathbb{C}\mathbf{1} \cong \mathbb{C}\mathbf{1}^\perp,$$

which is a special case of

$$\mathbb{E}/S \cong S^\perp$$

valid for any finite-dimensional complex vector space $\mathbb{E}$ with a Hermitian inner product, and any complex linear subspace $S \subset \mathbb{E}$. This isomorphism is a metric isomorphism. Here, $\mathbb{E}/S$ inherits a Hermitian inner product whose distance is given by Equation (21) with the group G replaced by S acting on $\mathbb{E}$ by translation and with the elements $\mathbf{q}_i$ in that formula being elements of $\mathbb{E}$. In the isomorphism, the metric we use on $S^\perp$ is the restriction of the metric from $\mathbb{E}$.

In our situation, S is the span of $\mathbf{1}$. Define

$$\mathbb{C}^2_0 := \mathbf{1}^\perp = \{\mathbf{q} : m_1q_1 + m_2q_2 + m_3q_3 = 0\}$$

to be the set of planar three-body configurations whose center of mass is at the origin. This two-dimensional complex space represents the quotient space of $\mathbb{C}^3$ by translations.

Jacobi coordinates: Diagonalizing the mass metric. We will need coordinates that diagonalize the restriction of the mass metric to $\mathbb{C}_0^2$. The quadratic form associated to the mass metric is the moment of inertia $I = \langle \mathbf{q}, \mathbf{q} \rangle$ of Equation (10). Thus, we look for coordinates Z_1, Z_2 on $\mathbb{C}_0^2$ such that

$$I = |Z_1|^2 + |Z_2|^2 \quad \text{whenever } \mathbf{q} \perp \mathbf{1}. \tag{22}$$

These coordinates are traditionally attributed to Jacobi, despite having been found earlier and explained more clearly by Lagrange [**16**], p. 292.

Exercise 4. Show that $\mathbf{1} = (1, 1, 1)$, $E_1 = (\frac{1}{m_1}, -\frac{1}{m_2}, 0)$, and $E_2 = (\frac{-1}{m_1+m_2}, \frac{-1}{m_1+m_2}, \frac{1}{m_3})$ form an orthogonal (not orthonormal) basis relative to the mass metric on $\mathbb{C}^3$.

The corresponding coordinates $\langle \mathbf{q}, \mathbf{1} \rangle$, $\langle \mathbf{q}, E_1 \rangle$, $\langle \mathbf{q}, E_2 \rangle$ are orthogonal coordinates for $\mathbb{C}^3$.

Definition. The coordinates $\langle \mathbf{q}, E_1 \rangle = q_1 - q_2 := Q_{12}$ and $\langle \mathbf{q}, E_2 \rangle = q_3 - \frac{m_1 q_1 + m_2 q_2}{m_1 + m_2}$ are *Jacobi coordinates* for $\mathbb{C}_0^2 := \{\mathbf{q} \in \mathbb{C}^3 : \mathbf{q}_{cm} = 0\}$.

Jacobi coordinates are indicated in Figure 4.

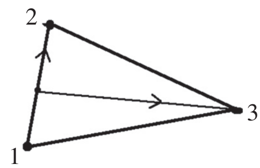

Figure 4. Jacobi vectors.

Normalizing the Jacobi coordinates yields our desired unitary diagonalizing coordinates $Z_i = \langle \mathbf{q}, e_i \rangle$, $i = 1, 2$ for $\mathbb{C}_0^2$ where $e_i = E_i / \|E_i\|$. We compute

$$Z_1 = \mu_1 (q_1 - q_2), \quad Z_2 = \mu_2 \left(q_3 - \frac{m_1 q_1 + m_2 q_2}{m_1 + m_2} \right) \tag{23}$$

with $\frac{1}{\mu_1^2} = \|E_1\|^2 = \frac{1}{m_1} + \frac{1}{m_2}$ and $\frac{1}{\mu_2^2} = \|E_2\|^2 = \frac{1}{m_3} + \frac{1}{m_1+m_2}$. These normalized Jacobi coordinates define the complex linear projection

$$\pi_{tr} : \mathbb{C}^3 \to \mathbb{C}^2 \quad \pi_{tr}(q_1, q_2, q_3) = (Z_1, Z_2), \tag{24}$$

which realizes the metric quotient of $\mathbb{C}^3$ by translations.

Dividing by rotations. It remains to divide $\mathbb{C}_0^2$ by the action of the rotation group S^1. A rotation by θ radians acts on the triangle's vertices q_j by $q_j \mapsto e^{i\theta} q_j$ and so it acts on the normalized Jacobi coordinates by $(Z_1, Z_2) \mapsto (e^{i\theta} Z_1, e^{i\theta} Z_2)$. We want to understand the resulting equivalence classes under rotation.

The functions $Z_i \bar{Z}_j$, $i, j = 1, 2$ are invariant under rotation. Put these functions into a 2 by 2 Hermitian matrix

$$\Phi(Z_1, Z_2) = \begin{pmatrix} |Z_1|^2 & Z_1 \bar{Z}_2 \\ \bar{Z}_1 Z_2 & |Z_2|^2 \end{pmatrix} = A, \tag{25}$$

or

$$\Phi(\mathbf{Z}) = \mathbf{Z}^T \bar{\mathbf{Z}} \tag{26}$$

where $\mathbf{Z} = (Z_1, Z_2)$,

$$\mathbf{Z}^T = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \quad \text{and} \quad \bar{\mathbf{Z}} = (\bar{Z}_1, \bar{Z}_2).$$

From the factorization (26), we see that $\Phi(\mathbf{Z})\mathbf{Z}^T = (|Z_1|^2 + |Z_2|^2)\mathbf{Z}^T$ while $\Phi(\mathbf{Z})\mathbf{W}^T = 0$ for $\mathbf{W} \perp \mathbf{Z}$. Thus, $\Phi(Z_1, Z_2)$ is the matrix of orthogonal projection onto the complex line spanned by $\mathbf{Z}$, multiplied by $\|\mathbf{Z}\|^2$. Now, two nonzero vectors $\mathbf{Z}$, $\mathbf{U}$ are related by rotation if and only if they span the same complex line and their lengths are equal. It follows that the image of Φ represents the quotient space $\mathbb{C}^2/S^1$, and Φ can be considered as the quotient map. We have just seen that the image of Φ consists of the Hermitian matrices of rank 1 whose nonzero eigenvalue is positive (corresponding to $\|\mathbf{Z}\|^2$), together with the zero matrix (corresponding to $\mathbf{Z} = 0$). In terms of the determinant and trace, these conditions on A are $\det(A) = 0$ and $tr(A) \geq 0$. Coordinatize Hermitian matrices as

$$A = \begin{pmatrix} w_4 + w_1 & w_2 + i w_3 \\ w_2 - i w_3 & w_4 - w_1 \end{pmatrix} \quad \text{for } w_j \text{ real}, \tag{27}$$

so that $\det(A) = w_4^2 - w_1^2 - w_2^2 - w_3^2$ and $tr(A) = w_4$. The discussion we have just had proves the following.

Proposition 2. *The image of the map Φ is the cone of 2 by 2 Hermitian matrices A (Equation (27)) satisfying*

$$w_4^2 - w_1^2 - w_2^2 - w_3^2 = 0 \tag{28}$$

and

$$w_4 \geq 0. \tag{29}$$

This cone realizes the quotient $\mathbb{C}^2/S^1$ with Φ implementing the quotient map $\mathbb{C}^2 \to \mathbb{C}^2/S^1$.

Now, map the real four-dimensional space of Hermitian matrices onto $\mathbb{R}^3$ by projecting out the trace w_4

$$(w_1, w_2, w_3, w_4) \mapsto pr(w_1, w_2, w_3, w_4) = (w_1, w_2, w_3).$$

The restriction of this linear projection to the cone given by the Equations (28) and (29) is a homeomorphism onto $\mathbb{R}^3$. Indeed solve the cone equations for w_4 to find $w_4 = +\sqrt{w_1^2 + w_2^2 + w_3^2}$, and hence, the inverse of the restricted projection is

$$(w_1, w_2, w_3) \mapsto (w_1, w_2, w_3, \sqrt{w_1^2 + w_2^2 + w_3^2}).$$

We have proved the following.

Proposition 3. *The map*

$$\pi^{rot} = pr \circ \Phi : \mathbb{C}^2 \to \mathbb{R}^3, \tag{30}$$

given by

$$\pi^{rot}(Z_1, Z_2) = \left(\frac{1}{2}(|Z_1|^2 - |Z_2|^2), Re(Z_1\bar{Z}_2), Im(Z_1\bar{Z}_2)\right) = (w_1, w_2, w_3), \tag{31}$$

realizes $\mathbb{R}^3$ *as the quotient space of* $\mathbb{C}^2$ *by the rotation group* S^1.

Remark. The restriction of the map (30) to the sphere $w_4 = 1$ is the famous *Hopf map* from the three sphere to the two sphere.

Proof of Theorem 1. We form $\pi = \pi^{rot} \circ \pi_{tr}$ by composing the linear projection π_{tr} of Equation (24) with the map π^{rot} immediately above. The first map realizes the quotient by translations and the second realizes the quotient by rotations, so together they realize the full quotient by the group of rigid motions. This establishes parts (A) and (B) of the theorem. Part (C), that the only triangles sent to $0 \in \mathbb{R}^3$ are the triple collision triangles $\mathbf{q} = (q, q, q)$, follows directly from the formulae for π^{rot} and π_{tr}. Indeed, the only point of $\mathbb{C}^2$ mapped to $0 \in \mathbb{R}^3$ by π^{rot} is the origin 0, and the only points of $\mathbb{C}^3$ mapped to $0 \in \mathbb{C}^2$ by π_{tr} are the triple collision triangles.

We verify the second half of part (D), which says w_3 is a constant times the oriented area of the triangle. We have $w_3 = -Z_1 \wedge Z_2$. The wedge of Equation (14) represents the oriented area of the parallelogram whose edges are $z = x + iy$ and $w = u + iv$. Thus, the oriented area of our triangle is $\frac{1}{2}(Q_{21}) \wedge (Q_{31})$ where we write $Q_{ij} = q_i - q_j$ for the edge connecting vertex j to vertex i. We have $Z_1 = \mu_1 Q_{12}$ and $Z_2 = \mu_2(p_1 Q_{31} + p_2 Q_{32})$ where $p_1 = m_1/(m_1 + m_2)$ and $p_2 = m_2/(m_1 + m_2)$ so that $p_1 + p_2 = 1$. Use $Q_{12} + Q_{23} + Q_{31} = 0$ and $Q_{ij} = -Q_{ji}$ to compute that $Z_2 = \mu_2(Q_{31} - p_2 Q_{12})$. Now, the wedge operation is skew symmetric: $Q_{12} \wedge Q_{12} = 0$. It follows that $w_3 = -\mu_1\mu_2\frac{1}{2}Q_{12} \wedge Q_{31} = +\mu_1\mu_2\frac{1}{2}Q_{21} \wedge Q_{31}$ as desired. The first half of part (D) follows immediately from this formula for w_3.

To establish part (E) regarding the operation of reflection on triangles, observe that we can reflect triangle $\mathbf{q}$ by changing all vertices q_i to $\bar{q}_i$, which in turn changes (Z_1, Z_2) to its conjugate vector $(\bar{Z}_1, \bar{Z}_2)$. This conjugation operation leaves w_1 and w_2 unchanged and changes w_3 to $-w_3$; the oriented area flips sign.

Part (F) is a computation. Observe from the Equations (25, 27) that $w_4 = \frac{1}{2}I$ and recall the cone condition Equation (28): $w_4^2 = w_1^2 + w_2^2 + w_3^2$. For part (G), see the paragraph immediately following Theorem 3 below. ■

8. MECHANICS VIA LAGRANGIANS. One of our goals is to write out reduced equations that encode Newton's Equations (1) as ODEs on shape space. The strategy for achieving this goal is to push the least action principle for the three-body problem down from the space $\mathbb{C}^3$ of located triangles to our shape space $\mathbb{R}^3$. We begin by stating the least action principle.

A classical mechanical system can be encoded by its *Lagrangian L* ([**1**] or [**8**]),

$$L = K - V, \tag{32}$$

the difference of its kinetic (K) and potential (V) energies. (The energy is the sum $K + V$.) Integrating the Lagrangian over a path c in the configuration space Q of the mechanical system defines that path's *action*

$$A[c] = \int_c L dt = \int_a^b L(c(t), \dot{c}(t)) dt.$$

In the last expression, the time interval $[a, b]$ parameterizes c so that $c : [a, b] \to Q$. The *principle of least action* asserts that the curve c satisfies Newton's equations if and only if c minimizes A among all paths $\gamma : [a, b] \to Q$ for which $c(a) = \gamma(a)$ and $c(b) = \gamma(b)$.

The principle is not a theorem! It is a guiding principle. To turn the principle into a theorem for the three-body problem requires careful wording and more hypotheses.

Theorem 2. *If a curve* $c : [0, T] \to \mathbb{C}^3$ *is collision-free on the open interval* $(0, T)$ *and minimizes the action among all curves sharing its endpoints* $c(0)$ *and* $c(T)$*, then that curve solves Newton's equations on* $(0, T)$*. Conversely, if* c *satisfies Newton's equations, then* c *is a critical point (but not necessarily a minimizer) of the action functional restricted to all curves sharing its endpoints.*

Remark. When we say that "c is a critical point" of the action, we mean that the derivative of the action is zero at c. How do we take the derivative of a function with respect to a curve? The calculus of variations is a wide-ranging subject that makes sense of such derivatives. Its roots go back to Leibnitz, Euler, Maupertuis, and Lagrange. See [**12**] for some of the fascinating history of the subject.

A theorem analogous to theorem 2 holds in situations more general than the three-body problem. For example, we could take $Q = \mathbb{R}^n$, K the squared norm associated to any inner product on $\mathbb{R}^n$, and $V : \mathbb{R}^n \to \mathbb{R}$ any smooth function. Then the Lagrangian $L(x, v) = K(v) - V(x)$ is a function on the phase space $\mathbb{R}^n \times \mathbb{R}^n$. Newton's equations are the second order differential equation

$$\ddot{c} = -\nabla V(c) \tag{33}$$

for curves c on $\mathbb{R}^n$ and a principle of least action holds for them. In these equations, ∇ is the gradient associated to the kinetic energy K, as described in analogy to Equation (15). More generally, we could take Q to be any Riemannian manifold, K the associated kinetic energy $\frac{1}{2}\Sigma g_{ij}(q)v_i v_j = \frac{1}{2}\langle v, v\rangle_q$, and $V : Q \to \mathbb{R}$ any smooth function. Then the ∇ of Newton's equations above is the covariant derivative, and the second derivative of c becomes the second covariant derivative D^2c/dt^2 of the curve.

Euler–Lagrange equations. Let $\xi^a, a = 1, \ldots, n$ be coordinates on our configuration space Q. Then the Lagrangian is a function of the ξ^a and its formal time derivatives $\dot{\xi}^a$:

$$L = L(\xi^1, \ldots, \xi^n, \dot{\xi}^1, \ldots, \dot{\xi}^n).$$

The *Euler–Lagrange equations*,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\xi}^a}\right) = \frac{\partial L}{\partial \xi^a} \tag{34}$$

are ODEs that a path $\xi^a = \xi^a(t)$ must satisfy if it minimizes the action. They are Newton's equations expressed in the new coordinates ξ_a.

Words are in order regarding the left-hand side of the EL (or Euler–Lagrange) Equations (34). We compute $\frac{\partial L}{\partial \dot{\xi}^a}$ by treating ξ^a and $\dot{\xi}^a$ as independent variables. The resulting $\frac{\partial L}{\partial \dot{\xi}^a}$ is now a function of the variables ξ^a and $\dot{\xi}^a$. We then compute $\frac{d}{dt}(\frac{\partial L}{\partial \dot{\xi}^a})$ by formally replacing the independent variables ξ^a and $\dot{\xi}^a$ in $\frac{\partial L}{\partial \dot{\xi}^a}$ by an alleged curve $\xi^a(t)$ and its time derivatives $\dot{\xi}^a(t)$ so as to get a function of time that we finally differentiate formally using the chain rule.

Exercise 5. Suppose that $K = \frac{1}{2}\Sigma g_{ab}\dot{\xi}^a\dot{\xi}^b$ and that $V = V(\xi^1, \dots, \xi^n)$. Verify that Newton's Equations (16) are equivalent to the *Euler–Lagrange equations* with respect to the coordinates ξ^a.

One of the beauties of the action principle is it is coordinate independent. If a path minimizes the action, then it does not matter what coordinate system we use to express that path: The path still minimizes the action and so satisfies the Euler–Lagrange equations in that coordinate system.

Exercise 6. For $Q = \mathbb{R}^2$ and $L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2)$ the EL equations are ODEs whose solutions are straight lines travelled at constant speed. Rewrite L in polar coordinates r, θ and write down the corresponding Euler–Lagrange equations, thus deriving ODEs for straight lines in polar coordinates.

Reducing the least action principle. The curves competing in the least action principle are subject to boundary conditions. In the principle as we stated it, the curves connect two fixed points. Replace the fixed points by fixed oriented congruence classes to get new boundary conditions. If we remember that an oriented congruence class is represented by a point of shape space, we arrive at an action principle for shape space.

Shape space action principle. Fix two shapes $\mathbf{w}_0, \mathbf{w}_1$ in the shape space $\mathbb{R}^3$. Suppose that the curve $\mathbf{q}(t) \in \mathbb{C}^3$ for $0 \le t \le T$ minimizes the standard action (32) among all curves in the space $\mathbb{C}^3$ of located triangles that join the corresponding oriented congruence classes $\Sigma_0 = \pi^{-1}(\mathbf{w}_0)$, $\Sigma_1 = \pi^{-1}(\mathbf{w}_1) \subset \mathbb{C}^3$ in time T. Then we will say that its projected curve $\mathbf{w}(t) = \pi(\mathbf{q}(t)) \in \mathbb{R}^3$ minimizes the *shape space action* among all curves connecting the endpoints $\mathbf{w}_0, \mathbf{w}_1$ in time T.

Consider an analogous change of boundary conditions for the simplest action functional in the plane, the length functional. Instead of minimizing the length of curves among all curves connecting two fixed points in the plane, replace the two points by two concentric circles Σ_0 and Σ_1. We know that the minimizer will be a radial line segment, perpendicular to both Σ_0 and Σ_1 at its endpoints. More generally, for a Lagrangian on $\mathbb{R}^n$ of the general form (32) , if we replace the fixed endpoint minimization problem with the problem of minimizing the action among all curves connecting two given *subspaces* $\Sigma_0, \Sigma_1 \subset \mathbb{R}^n$, then we induce a derivative condition at the endpoints. Namely extremal curves, in addition to satisfying the Euler–Lagrange equations, must hit their targets orthogonally: $\dot{c}(0) \perp \Sigma_0$ at $c(0)$ and $\dot{c}(T) \perp \Sigma_1$ at $c(T)$. We call this added condition " first variation orthogonality."

Returning to our situation, $\Sigma_0 = \pi^{-1}(\mathbf{w}_0)$, $\Sigma_1 = \pi^{-1}(\mathbf{w}_1) \subset \mathbb{C}^3 = \mathbb{R}^6$. We will interpret first variation orthogonality in mechanical terms. Σ_0 is formed by applying variable rigid motions $g \in G$ to any single point $\mathbf{q}_0 \in \Sigma_0$. Let $g = g(t)$ be any smooth path in G, and form the corresponding path $g(t)\mathbf{q}_0$ in Σ_0. Differentiating this path and then alternately taking $g(t)$ to be a curve of translations or a curve of rotations, we see that the tangent space $T_{\mathbf{q}_0}\Sigma_0$ to Σ_0 at $\mathbf{q}_0$ is spanned by two subspaces, $\{\dot{c}\mathbf{1}, \dot{c} \in \mathbb{C}\}$ for translations and $\{i\dot{\theta}\mathbf{q}_0 : \dot{\theta} \in \mathbb{R}\}$

$$T_{\mathbf{q}_0}\Sigma_0 = \text{infinitesimal rigid motions} \tag{35}$$

$$= (\text{translational}) + (\text{rotational}) \tag{36}$$

$$= span_{\mathbb{C}}\mathbf{1} + span_{\mathbb{R}}(i\mathbf{q_0}). \tag{37}$$

The first variation orthogonality condition is thus the condition that our extremal $\mathbf{q}(t)$ be orthogonal to both the translation and rotational spaces: $\langle \mathbf{1}, \dot{\mathbf{q}}(0)\rangle = 0$ and $\langle i\dot{\theta}\mathbf{q}(0), \dot{\mathbf{q}}(0)\rangle_{\mathbb{R}} = 0$. But as we saw in the Equations (12, 11), these orthogonality conditions are equivalent to the assertions that the linear and angular momentum are zero at $\mathbf{q}_0$. (The inner product used for orthogonality is the real part of the Hermitian mass inner product and $Im(\langle \mathbf{q_0}, \dot{\mathbf{q}}\rangle) = Re(\langle i\mathbf{q_0}, \dot{\mathbf{q}}\rangle$.) We summarize as follows.

Lemma 1. *The curve $\mathbf{q}(t)$ in $\mathbb{C}^3$ is orthogonal to the oriented congruence class through $\mathbf{q}_0 = \mathbf{q}(0)$ if and only if its linear and angular momentum are zero at $t = 0$. (See the Equations (12) and (11).)*

Now, if the curve $\mathbf{q}(t)$ of this lemma is an extremal for our shape space action principle, then it must satisfy the EL equations that are Newton's equations. Since linear and angular momentum are conserved for solutions to Newton's equations, we have that the linear and angular momentum are identically zero all along the curve. Equivalently, if an extremal curve is orthogonal to the group orbit Σ_0 through one of its points $\mathbf{q}(0)$, then it is orthogonal to the group orbit Σ_t through every one of its points $\mathbf{q}(t)$. We have established the following.

Proposition 4. *The extremals for the shape space action principle are precisely those solutions to Newton's equations whose linear and angular momentum are zero.*

The proposition suggests a strategy for finding a Lagrangian L_{shape} on shape space whose action minimization is equivalent the shape space action principle. Break up kinetic energy into

$$K = \text{translational part} + \text{ rotational part } + \text{ shape part.} \tag{38}$$

We have just agreed that the translation and rotational part of the kinetic energy must be zero along our shape extremals corresponding to the fact that they are orthogonal to G-orbits. Let us denote the last term, the shape term of the kinetic energy, as K_{shape}. Thus,

$$L_{shape} = K_{shape} - V \tag{39}$$

is the shape Lagrangian. It remains to express K_{shape} in terms of shape coordinates w_i and their time derivatives $\dot{w}_i$ and V in terms of the w_i.

Shape kinetic energy. The decomposition (38) applied to velocities is sometimes called the "Saari decomposition" ([**22**], [**23**], [**3**] p.331).

$$\dot{\mathbf{q}} = (\text{translational part} + \text{ rotational part }) + \text{ shape part.} \tag{40}$$

$$= \quad T_q(Gq) \quad \oplus \quad (T_q(Gq))^{\perp} \tag{41}$$

$$= \quad \text{vertical} \quad \oplus \quad \text{horizontal.} \tag{42}$$

In the differential geometry of bundles, such a splitting of tangent vectors is known as a "'vertical-horizontal" splitting. The group directions $T_q(Gq)$ form the "vertical space." The orthogonal complement $T_q(Gq))^{\perp}$ to the vertical space forms the "horizontal space." This vertical-horizontal decomposition, which depends on the base point $\mathbf{q}$ at which velocities are attached, is orthogonal and leads to the following.

Proposition 5. *Suppose that the center of mass of our located triangle is zero. Then the Saari decomposition of kinetic energy, Equations (38), is*

$$K = \frac{1}{2}\frac{\|\mathbf{P}\|^2}{M} + \frac{1}{2}\frac{J^2}{I} + \frac{1}{2}\frac{\|\dot{\mathbf{w}}\|^2}{I}$$

where $\dot{\mathbf{w}} = \frac{d}{dt}\pi(\mathbf{q}(t))$, $P = P(\dot{\mathbf{q}})$, and $J = J(\mathbf{q}, \dot{\mathbf{q}})$ are the linear and angular momenta (Equation 11, 12). In particular,

$$K_{shape} = \frac{1}{2}\frac{\|\dot{\mathbf{w}}\|^2}{I} \quad \text{where } I = 2\sqrt{\|\mathbf{w}\|}. \tag{43}$$

Proof. A real basis for the two-dimensional translational part of the motion consists of $\mathbf{1}$ and $i\mathbf{1}$. A real basis for the one-dimensional rotational part is $i\mathbf{q}$. The rotational part is orthogonal to the translational part since the center of mass is $\frac{1}{M}\langle \mathbf{q}, \mathbf{1}\rangle$ and has been set equal to zero. Hence, $\mathbf{1}$, $i\mathbf{1}$, and $i\mathbf{q}$ is an orthogonal basis for the vertical part, $T_q(Gq)$. Normalize to get the real orthonormal basis

$$e_1, e_2, e_3 = \mathbf{1}/\sqrt{M}, i\mathbf{1}/\sqrt{M}, i\mathbf{q}/\sqrt{I}, \qquad \text{where } M = \langle \mathbf{1}, \mathbf{1}\rangle = m_1 + m_2 + m_3,$$

for the vertical space. Let $\dot{\mathbf{q}} \in \mathbb{C}^3$ be an arbitrary vector based at the located triangle $\mathbf{q} \in \mathbb{C}^3$. Expand this vector out as an orthogonal direct sum to get a quantitative form of the Saari decomposition Equations (42)

$$\dot{\mathbf{q}} = \langle \dot{\mathbf{q}}, e_1\rangle_{\mathbb{R}} e_1 + \langle \dot{\mathbf{q}}, e_2\rangle_{\mathbb{R}} e_2 + \langle \dot{\mathbf{q}}, e_3\rangle_{\mathbb{R}} e_3 + (\text{shape}).$$

The first three terms form the vertical part of $\dot{\mathbf{q}}$ in Equation (42) while the final (shape) part is, by definition, orthogonal to the first three terms and forms the horizontal part. Squaring lengths and using the orthonormality of e_1, e_2, e_3, we find that

$$\langle \dot{\mathbf{q}}, \dot{\mathbf{q}}\rangle = |P|^2/M + J^2/I + \text{shape}^2.$$

It remains to show that $|\text{shape}|^2 = \frac{\|\dot{\mathbf{w}}\|^2}{I}$. In other words, we need to show that

$$\|\dot{\mathbf{w}}\|^2 = \|\mathbf{q}\|^2\|\dot{\mathbf{q}}\|^2 \quad \text{if} \quad P(\dot{\mathbf{q}}) = 0, \quad J(\mathbf{q}, \dot{\mathbf{q}}) = 0 \quad \text{and} \quad \dot{\mathbf{w}} = D\pi_{\mathbf{q}}(\dot{\mathbf{q}}). \tag{44}$$

To this end, write out the map π^{rot} in real coordinates, using $Z_j = x_j + iy_j$, $\mathbf{Z} = (Z_1, Z_2) = (x_1, y_1, x_2, y_2)$. We have $\pi^{rot}(x_1, y_1, x_2, y_2) = (\frac{1}{2}(x_1^2 + y_1^2 - x_2^2 - y_2^2), x_1x_2 + y_1y_2, x_2y_1 - x_1y_2)$. Compute the Jacobian

$$D\pi^{rot}_{\mathbf{Z}} = \begin{pmatrix} x_1 & y_1 & -x_2 & -y_2 \\ x_2 & y_2 & x_1 & y_1 \\ -y_2 & x_2 & y_1 & -x_1 \end{pmatrix}.$$

Set

$$L = D\pi^{rot}_{\mathbf{Z}},\ D\pi_{\mathbf{q}} = L \circ \pi^{tr}.$$

In the last equality, π denotes the quotient map of Theorem 1, and we used the fact that $\pi = \pi^{rot} \circ \pi_{tr}$ with π_{tr} linear. The three rows of L are orthogonal and have length $\|\mathbf{Z}\|^2 = x_1^2 + y_1^2 + x_2^2 + y_2^2$. Now, $\mathbf{q} \in \mathbb{C}^2_0$ so that $\|\mathbf{Z}\|^2 = \|\mathbf{q}\|^2$ and π_{tr} is a linear isometry of $\mathbb{C}^2_0$ to $\mathbb{C}^2$. It follows that

$$LL^T = \|\mathbf{q}\|^2 Id = D\pi_{\mathbf{q}} D\pi^T_{\mathbf{q}}.$$

The kernel of $D\pi = D\pi_{\mathbf{q}}$ is the vertical space, the span of e_1, e_2, e_3, since π is invariant under rotations and translations. Thus, the image of $D\pi^T$ is the horizontal space, which is the orthogonal complement to e_1, e_2, e_3 and is the space called "(shape)" above. Consequently, any vector $\dot{\mathbf{q}}$ of the form required in Equation (44) can be written $\dot{\mathbf{q}} = D\pi^T \mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^3$. Thus,

$$\|\dot{\mathbf{q}}\|^2 = \langle D\pi^T \mathbf{v}, D\pi^T \mathbf{v}\rangle \tag{45}$$

$$= \langle \mathbf{v}, D\pi D\pi^T \mathbf{v}\rangle \tag{46}$$

$$= \langle \mathbf{v}, \|\mathbf{q}\|^2 \mathbf{v}\rangle \tag{47}$$

$$= \|\mathbf{q}\|^2 \|\mathbf{v}\|^2. \tag{48}$$

Moreover, $\dot{\mathbf{w}} = D\pi(\dot{\mathbf{q}})$ so that $\dot{\mathbf{w}} = D\pi D\pi^T(\mathbf{v}) = \|\mathbf{q}\|^2 \mathbf{v}$. We get that $\|\dot{\mathbf{w}}\|^2 = \|\mathbf{q}\|^4 \|\mathbf{v}\|^2 = \|\mathbf{q}\|^2 \|\dot{\mathbf{q}}\|^2$. Thus, $\|\dot{\mathbf{q}}\|^2 = \|\dot{\mathbf{w}}\|^2 / \|\mathbf{q}\|^2$. Finally, use $\|\mathbf{q}\|^2 = I$ and $I = 2\sqrt{\|\mathbf{w}\|}$ ((F) of Theorem 1). ■

9. SHAPE SPACE METRIC.

Definition (Shape space metric). The shape space metric is twice the shape space kinetic energy K_{shape}, when viewed as a Riemannian metric on shape space.

We saw in the previous proposition that the shape space metric is given by

$$ds^2_{shape} = \frac{dw_1^2 + dw_2^2 + dw_3^2}{2\sqrt{w_1^2 + w_2^2 + w_3^2}}. \tag{49}$$

Define the *length* ℓ of a path c in shape space to be $\ell(c) = \int_c ds_{shape} := \int_a^b \sqrt{2K_{shape}}dt$. Define the distance between two points of shape space to be the infinum of the lengths of all paths joining the two points. In other words, the shape space length is the action relative to the Lagrangian $\sqrt{2K_{shape}}$, and shape space distance between two points is realized by a length-minimizing curve joining them. We call such a length-minimizing curve a *geodesic*.

Recall the Cauchy–Schwartz inequality. $\int f(t)g(t)dt \leq \sqrt{\int f(t)^2 dt}\sqrt{\int g(t)^2 dt}$ with equality if and only if $f(t) = cg(t)$ (a.e.), with c a constant. Applied to $f = 1, g(t) = \sqrt{2K_{shape}(t)}$, we get $\ell(\gamma) \leq \sqrt{b-a}\sqrt{2\int K_{shape}dt}$ with equality if and only if K_{shape} is constant along the curve γ. Reparameterizing a curve γ does not change its length, and we can always parameterize γ so that its square speed $2K_{shape}$ is constant. It follows from Cauchy–Schwartz that the curves that minimize $\int_\gamma K_{shape}dt$ are precisely the constant speed geodesics. The shape space action principle holds just as well for K in place of $K - V$, and the reduced Lagrangian for K is K_{shape}. The geodesics for K are straight lines in $\mathbb{C}^3$. Putting together these observations, we have proved the assertions of the first two sentences of the next theorem.

Theorem 3. *The distance function defined by the shape space metric agrees with the shape space distance of Equation (21) . Its geodesics are the projections by $\pi : \mathbb{C}^3 \to \mathbb{R}^3$ of horizontal lines in $\mathbb{C}^3$. Each plane $\Pi : Aw_1 + Bw_2 + Cw_3 = 0$ through the origin is totally geodesic: A geodesic that starts on Π initially tangent to Π, lies completely in the plane Π. The restriction of the Riemannian metric to such a plane Π, when expressed in standard Euclidean polar coordinates (r, θ) on that plane, has the form*

$$ds^2_{shape}|_\Pi = dr^2 + \frac{1}{4}r^2 d\theta^2.$$

In order to finish the proof of this theorem, let $\ell(t) = \mathbf{q} + t\mathbf{v}$ be a horizontal line passing through the point $\mathbf{q} \in \mathbb{C}^2_0 \subset \mathbb{C}^3$, $\mathbf{q} \neq 0$ with horizontal tangent vector $\mathbf{v}$. There are two possibilities for $\mathbf{v}$: Either $\mathbf{v}$ is a multiple of $\mathbf{q}$ or $\mathbf{v}$ is linearly independent from $\mathbf{q}$. In the first case, we may assume that $\mathbf{v} = \mathbf{q}$ is the radial vector. Then ℓ is a radial line and $\pi(\ell)$ is the ray connecting the triple collision point 0 to $\mathbf{w} = \pi(\mathbf{q})$ (traversed twice). The distance from 0 to $\mathbf{w}$ along this ray is the radial variable

$$r = dist(0, \mathbf{w}) = \|\mathbf{q}\| = \sqrt{I} = \sqrt{2\|\mathbf{w}\|}.$$

In the second case, $\mathbf{q}$ and $\mathbf{v}$ span a real horizontal two-plane P in $\mathbb{C}^3$ that passes through 0 and contains the line ℓ. One computes that the projection $\Pi := \pi(P) \subset \mathbb{R}^3$ is a plane (relative to the coordinates w_i) passing through 0. However, the projection $\pi(\ell)$ is not a line (relative to the linear coordinates w_i)!

We can understand the geodesic $\pi(\ell)$ in shape space by understanding the restriction $ds^2_{shape}|_\Pi$ of the shape space metric to the plane $\pi(P)$. Here is what we know so far about this metric. Radial lines are geodesics. The distance along a radial line from triple collision point 0 to $\mathbf{w} \in \pi(P)$ is r as given above. To dilate the metric by a factor $t > 0$, we multiply $\mathbf{w} \in \mathbb{R}^3$ by t^2 since $\mathbf{q} \mapsto t\mathbf{q}$ corresponds to $\mathbf{w} \mapsto t^2\mathbf{w}$ under π. Finally, the metric on Π is rotationally symmetric since the expression (49) is rotationally invariant. From all of this information, we deduce that the restricted metric has the form

$$ds^2_{\pi(P)} = dr^2 + c^2 r^2 d\theta^2, \tag{50}$$

where (r, θ) are polar coordinates on the plane and c is a constant. It remains to show that $c = 1/2$. With this in mind, consider the circle $r = 1$ in the plane Π. Its circumference computed from Equation (50) is $2\pi c$. But we can also compute its length by working up on $P \subset \mathbb{C}^3$. Take $\mathbf{q}$ and $\mathbf{v}$ to both be unit length and orthogonal so that they form an orthonormal basis for P. Then the corresponding horizontal circle on P is $\cos(s)\mathbf{q} + \sin(s)\mathbf{v}$, for $0 \leq s \leq 2\pi$. But $\pi(\mathbf{q}) = \pi(-\mathbf{q})$ since $-\mathbf{q} = e^{i\pi}\mathbf{q}$ so that the π-projection of this circle closes up once we have gone half way around, from $s = 0$ to $s = \pi$. Thus, the projected circle on $\pi(P) = \Pi$ has the length of half a unit circle, or π. Comparing lengths, we see that $c = 1/2$.

Any metric of form (50) is that of a cone. We can form our $c = 1/2$ cone by taking a sheet of paper and marking the midpoint of one edge to be the cone point (see Figure 5). Fold that edge up so the two halves touch each other, and we have a paper model of the required cone. Note that the circle of radius r about the cone point has circumference πr, as required.

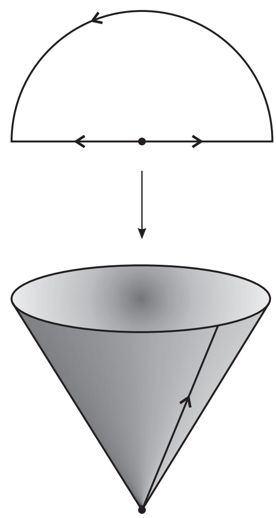

Figure 5. Folding a half-sheet of paper makes the desired cone.

10. POTENTIAL ON SHAPE SPACE. We need a formula for the three triangle side lengths r_{ij} in terms of the shape coordinates w_i's in order to express the potential (Equation (8)) as a function on shape space. Let $\mathbf{b}_{ij}$ be the point on the shape sphere $\{\|\mathbf{w}\| = 1\}$, which represents the ij collision: $r_{ij} = 0$. The desired equation is

$$r_{ij}^2 = \frac{m_i + m_j}{m_i m_j}(\|\mathbf{w}\| - \mathbf{w} \cdot \mathbf{b}_{ij}). \tag{51}$$

An important geometric fact underlies this computation. Let $d_{ij}(\mathbf{w}) = d(\mathbf{w}, \mathbb{R}^+\mathbf{b}_{ij})$ denote the shape space distance (Equation (21)) from the shape space point $\mathbf{w}$ to the ij binary collision ray $\mathbb{R}^+\mathbf{b}_{ij}$. Then

$$d_{ij} = \sqrt{\mu_{ij}} r_{ij} \tag{52}$$

where $\mu_{ij} = m_i m_j/(m_i + m_j)$. For proofs of this formula or Equation (51), see [**5**], [**18**].

The potential is homogeneous of degree -1 so that

$$V(r, \theta, \phi) = \frac{1}{r}\tilde{V}(\theta, \phi)$$

where $\tilde{V}$ is the restriction of V to the unit sphere $r = 1$ and ϕ, θ are standard spherical coordinates: $\mathbf{w} = \frac{r^2}{2}(\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi))$. A contour plot of V is indicated in Figure 7.

11. REDUCED EQUATIONS OF MOTION. Having written the shape space kinetic energy (Equation (43)) and potential energy (Equations (8), (51)) in terms of shape variables w_i, we have the shape space Lagrangian

$$L_{shape} = \frac{1}{2}\frac{\dot{w}_1^2 + \dot{w}_2^2 + \dot{w}_3^2}{2\sqrt{w_1^2 + w_2^2 + w_3^2}} + \frac{c_{12}}{r_{12}} + \frac{c_{23}}{r_{23}} + \frac{c_{13}}{r_{13}} \tag{53}$$

with constants $c_{ij} = (m_i m_j)^{3/2}/\sqrt{m_i + m_j}$. Its Euler–Lagrange Equations (34)

$$\frac{d}{dt}\left(\frac{\partial L_{shape}}{\partial \dot{w}_i}\right) = \frac{\partial L_{shape}}{\partial w_i} \quad \text{for } i = 1, 2, 3 \tag{54}$$

are our desired reduced equations and are the ODEs for the zero-angular momentum three-body equations as written in shape space.

12. INFINITELY MANY SYZYGIES. The shape space Lagrangian (Equation (53)), combined with the realization (Equation (52), is that of a point mass moving in $\mathbb{R}^3$, endowed with metric (49), subject to the attractive force generated by the pull of the three binary collision rays. These rays lie in the collinear plane $w_3 = 0$. Consequently, the point is always attracted toward the collinear plane. This physical analogy suggests that the point must oscillate back and forth, crossing that plane infinitely often.

Theorem 4 (See [18]). *If a solution with negative energy and zero angular momentum does not begin and end in triple collision, then it must cross the collinear plane $w_3 = 0$ infinitely often.*

Sketch of proof of theorem 4. The physical analogy just described led us to discover a differential equation of the form $\frac{d}{dt}(f\frac{d}{dt}z) = -gz$ for a normalized height variable $z = w_3/\tilde{I}$, where $\tilde{I} = r_{12}^2 + r_{23}^2 + r_{31}^2$. Here, f is a positive function on shape space and g is a non-negative function of the w_i and $\dot{w}_i$ that is positive away from the Lagrange homothety solution. (This solution is the special case of the Lagrange solution (Equation 18) in which $\lambda(t)$ is real. The solution evolves by scaling, ending in triple collision.) The theorem follows from the differential equation by a Sturm–Liouville argument and the fact that near-triple collision behavior is governed by behavior near the Lagrange homothety solution.

13. FINALE. We end with another theorem whose conception and proof relies on the shape space reformulation of the three-body problem.

Theorem 5 (See [5]). *There is a periodic solution to the equal mass zero angular momentum three-body problem in which all three masses chase each other around the same figure-eight-shaped curve.*

Sketch of proof of Theorem 5. There are three types of isosceles triangles, depending on which mass forms the vertex. Each type defines a "longitude"—a great circle on the shape sphere passing through the north and south poles. (See Figure 6.) Recall these poles represent the equilateral triangles, or Lagrange points. Each longitude intersects the equator of collinear triangles at two antipodal points, one of which is a binary collision point. In the equal mass case, the other point of intersection is an Euler point, the binary collision points and Euler points are equally spaced along the equator, and the three isosceles great circles together with the equator divide the sphere up into 12 congruent spherical triangles with angles 90-90-60. One key to the proof is that this apparent 12-fold symmetry is indeed a symmetry group of the differential equations.

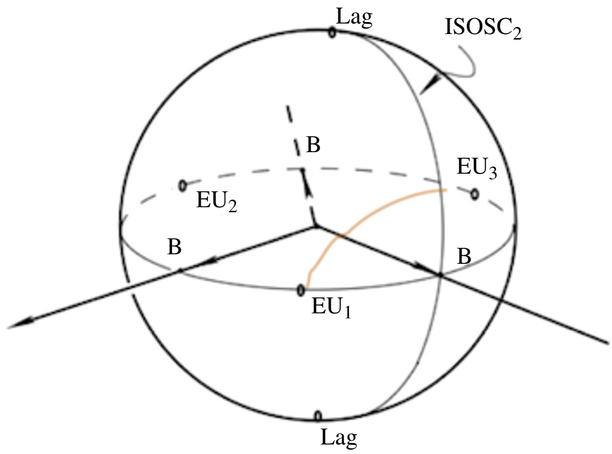

Figure 6. The shape sphere. The three binary collision rays pierce the shape sphere at the three points (B) that lie on the collinear equator that in turn also contains the three Euler points (EU). One isosceles great circle (*ISOSC*$_2$) is drawn. Like all isosceles circles, it joins the two Lagrange points (Lag), and passes through a binary collision point. The minimizer out of which the eight is built is drawn in a brown and connects an Euler point to *ISOSC*$_2$.

Viewed in shape space, instead of on the shape sphere, each isosceles great circle represents a plane of isosceles triangles, and each Euler point represents a ray of Euler collinear triangles. Label these planes $ISOSC_i$ for $i = 1, 2, 3$ and label these rays EU_i for $i = 1, 2, 3$. For example, $ISOSC_2$ is defined by the equation $r_{12} = r_{23}$. Consider the problem of minimizing the shape space action $\int_c L_{shape} dt$ among all paths $c : [0, T] \to \mathbb{R}^3$ connecting EU_1 to $ISOSC_2$. Suppose such a minimizer, call it γ_*, exists and is collision-free. First, variation orthogonality implies that γ_* must hit both the Euler ray and the isosceles plane orthogonally. (See Figure 6.) The minimizer γ_* will form one-twelfth of the figure-eight solution.

To build the rest of the eight from γ_*, we use equality of masses and the consequent 12-fold symmetry group. This order 12 group is generated by reflections about the isosceles planes and the equator, and each of its element are symmetries of the kinetic energy, the potential energy, the shape space Lagrangian, and consequently, of the reduced equations. Reflection R_E about the equator is induced by reflecting all three masses about any fixed line in the inertial plane and is a symmetry no matter what the masses are. The half-twist σ_{13} about the binary collision ray $r_{13} = 0$ is induced by the operation $(q_1, q_2, q_3) \mapsto (q_3, q_2, q_1)$ of interchanging masses 1 and 3 and is a symmetry provided $m_1 = m_3$. Reflection R_2 about $ISOSC_2$ is the composition of R_E with σ_{13}. Applying the symmetries in turn to our minimizer γ_*, we obtain 12 congruent curves that fit together continuously to "wrap" twice around the sphere. Due to first-variation orthogonality, they also fit together smoothly! For example, γ_* and its reflection $R_2(\gamma_*)$ about $ISOSC_2$ share the same derivative since both are orthogonal to the isosceles plane. In other words, $R\gamma_*(2T - t)$ is a reduced solution to Newton's equations whose shape space point and shape space velocity agree with those of γ_* at $t = T$. By unique dependence of solutions to ODEs on their initial conditions it follows that in order to continue the solution γ_* past $t = T$ through the isosceles plane, we simply concatenate $\gamma_*(t)$ with its reflected image $R\gamma_*(2T - t)$. Continuing in this manner with reflections or half-twists, we see see that the concatenation of these 12 congruent arcs, taken together, forms one smooth periodic solution to the reduced Newton's equation. The horizontal lift to $\mathbb{C}^3$ of this periodic solution is automatically a solution to Newton's equations. With extra work, we can show that this horizontal lift is itself periodic of period $12T$ (i.e., the solution is periodic, not just periodic modulo rigid motions) and is also a *choreography*.

Definition. An N-body choreography of period T is a solution to the N-body problem that has the particular form $\mathbf{q}(t) = (q_1(t), q_2(t), \ldots, q_N(t))$ where $q_i(t) = s(t - iT/N)$ for $i = 2, 3, \ldots, N$ for some fixed T periodic curve $s(t)$ in the plane (or space).

This curve is the figure eight. See Figure 7 for the curve γ, the eight projected to shape space.

We have skipped the difficult part of the proof, which is showing that the minimum γ_* exists and is collision-free. For this, we refer the reader to [**5**].

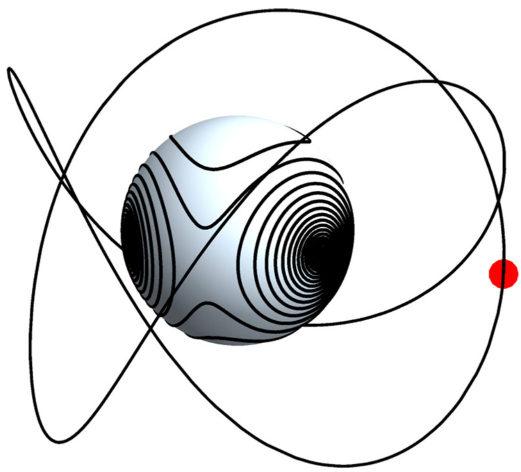

Figure 7. The figure eight as it appears in shape space

History. The figure-eight solution curve was discovered numerically in 1993 using the principle of least action by C. Moore [**20**] whose short article we highly recommend. Chenciner and myself rediscovered the solution using the shape space least action principle in 2000 [**5**]. Our methods yielded a rigorous existence proof and soon led to the discovery of a multitude of N-body choreographies. See, for example, [**7**], [**2**], [**4**], [**10**], and [**19**].

ACKNOWLEDGMENT. Deep thanks to Alain Chenciner, Alain Albouy, Carles Simó and Rick Moeckel for conversations and work together over the years. Additional thanks to Rick for Figures 7 and 6, and to the mysterious "The Logging Road Cyclist" (http://theloggingroadcyclist.com) for a careful reading and an outside perspective. The author thankfully acknowledges support of NSF grant DMS-20030177.

REFERENCES

1. V. I. Arnol'd, *Mathematical Methods in Celestial Mechanics*. Translated by A. Weinstein and K. Vogtman. Springer, Berlin, 1978.

2. K. C. Chen, Existence and minimizing properties of retrograde orbits to the three-body problem with various choices of masses, *Ann. of Math.* **167** (2008) 325–348, http://dx.doi.org/10.4007/annals.2008.167.325.
3. A. Chenciner, A l'infini en temps fini, *Séminaire Bourbaki* **832** (1997) 323–353.
4. A. Chenciner, J. Gerver, R. Montgomery, C. Simó, Simple choreographies of N bodies: A preliminary study, in *Geometry, Mechanics, and Dynamics*. Edited by P. Newton, P. Holmes. and A. Weinstein, Springer, Berlin, 2002, http://dx.doi.org/10.1007/0-387-21791-6_9.
5. A. Chenciner, R. Montgomery, A remarkable periodic solution of the three-body problem in the case of equal masses, *Ann. of Math.***152** (2000) 881–901, http://dx.doi.org/10.2307/2661357.
6. L. Euler, Considérations sur le problème de trois corps, *Mémoires de l'Académie de Sciences de Berlin* **19** (1770) 194–220.
7. D. Ferrario, S. Terracini, On the Existence of collisionless equivariant minimizers for the classical n-body problem, *Invent. Math.* **155** (2004) 305–362, http://dx.doi.org/10.1007/s00222-003-0322-7.
8. R. Feynman, R. Leighton, M. Sands, *The Feynman Lectures on Physics* **1**. Addison-Wesley, San Francisco, 1963.
9. T. Fujiwara, H. Fukuda, A. Kameyama, H. Ozaki and M.Yamada, Synchronised similar triangles, *J. Phys. A: Math. Gen.* **37** (2004) 10571–10584.
10. G. Fusco, G. F. Gronchi, P. Negrini, Platonic polyhedra, topological constraints and periodic solutions of the classical N-body problem, *Invent. Math.* **185** (2011) 283–332.
11. M. Herman, Some open problems in dynamical systems, *Proc. Int. Congress of Math.*, Documenta Mathematica J. DMV, Extra volume ICM II, (1998) 797–808.
12. S. Hildebrandt, A. Tromba *Mathematics and Optimal Form*. Scientific American Library, New York, 1985.
13. T. Iwai, A gauge theory for the quantum planar three-body problem. *J. Math. Phys.* **26** (1987) 964–974, http://dx.doi.org/10.1063/1.527588.
14. J. Lagrange, *Méchanique analitique*. First edition. Veuve Desaint, Paris, 1788.
15. D. G. Kendall, Shape manifolds, procrustean metrics, and complex projective space, *Bull. London Math. Soc.* **16** (1984) 81–121, http://dx.doi.org/10.1112/blms/16.2.81.
16. J.-L. Lagrange, Essai sur le Probléme des Trois Corps, in *Prix de l'Académie Royale des Sciences de Paris*. **IX**, 1772.
17. R. Moeckel, Chaotic dynamics near triple collision, *Arch. Rational Mech. Anal.* **107** (1989) 37–69, http://dx.doi.org/10.1007/bf00251426.
18. R. Montgomery, Infinitely many syzygies, *Arch. Rational Mech. Anal.* **164** (2002) 311–340, http://dx.doi.org/10.1007/s00205-002-0211-z.
19. J. Montaldi, K. Steckles, Classification of symmetry groups for planar n-body choreographies, *Forum Math. Sigma* **1** (2013), http://dx.doi.org/10.1017/fms.2013.5.
20. C. Moore, Braids in classical gravity, *Phys. Rev. Lett.* **70** (1993) 3675–3679.
21. I. Newton, *Philosophiæ Naturalis Principia Mathematica*. (1667), http://en.wikisource.org/wiki/The_Mathematical_Principles_of_Natural_Philosophy_%281846%29.
22. D. Saari, The manifold structure for collision and hyperbolic-parabolic orbits in the n-body problem, *J. Differential Equations* **55** (1984) 300–329, http://dx.doi.org/10.1016/0022-0396(84)90072-x.
23. ———, *Collisions,Rings,and Other Newtonian N-Body Problems*. CBMS **104**, AMS, Providence, R.I., 2005.

RICHARD MONTGOMERY did poorly in calculus at Berkeley High School as a senior in 1974. He taught whitewater kayaking professionally and engaged in first descents of rivers in California's Sierra Nevada and Coast Range from 1975 to 1985. He attended Columbia Junior College in the Sierra Nevada Foothills (1976-78) and Sonoma State University (1978–1980) just south of the California wine region where he met his wife Judith. He received his PhD under Jerry Marsden at U.C. Berkeley in 1985 and was sure that he would never work in celestial mechanics. A falling cat dragged him in to that work in 1996.
Mathematics Dept., U.C. Santa Cruz, Santa Cruz, CA, 95064
rmont@ucsc.edu