The geometric phase of the three-body problem

1996 Nonlinearity 9 1341

(http://iopscience.iop.org/0951-7715/9/5/014)

Nonlinearity **9** (1996) 1341–1360. Printed in the UK

The geometric phase of the three-body problem

Richard Montgomery†
Department of Mathematics, University of California, Santa Cruz, CA 95064, USA

Received 29 September 1995, in final form 20 May 1996
Recommended by J Laskar

Abstract. Suppose that the initial triangle formed by the three moving masses of the three-body problem is similar to the triangle formed at some later time. We derive a simple integral formula for the overall rotation relating the two triangles. The formula is based on the fact that the space of similarity classes of triangles forms a two-sphere which we call the shape sphere. The formula consists of a 'dynamic' and 'geometric' term. The geometric term is the integral of a universal two-form on a'reduced configuration space'. This space is a two-sphere bundle over the shape sphere. The fibring spheres are instantaneous versions of the angular momentum sphere appearing in rigid body motion. Our derivation of the formula is similar in spirit to our earlier reconstruction formula for the rigid body motion.

AMS classification scheme numbers: 70F07, 58F05

1. Introduction and results

1.1. A reconstruction formula

The three-body problem concerns understanding the motions of three point masses travelling in space according to Newton's laws of mechanics. The three masses form a triangle in space so that Newton's equations define a dynamical system on the space of triangles. The shape (congruence class) of the triangle is the primary variable. Shape variables are further divided up into an overall scale parameter I, and the similarity class of the triangle. The similarity classes form a two-sphere, denoted S, and called the shape sphere. We view the orientation and position of the triangle in space as secondary variables. (The translational part of the motion is eliminated by the usual trick of going to centre-of-mass coordinates.) Our basic question is: given that the initial and final triangles of a three-body motion are similar, what is the rotation, R, which relates the two triangles (up to scale)?

We will suppose that the planes defined by the initial and final triangles and the total angular momentum vector, $\mathbf{J}_0$ are known. Write $\mathbf{J}_0 = J_0\mathbf{e}_3$ where J_0 is the magnitude of the angular momentum and $\mathbf{e}_3$ is a unit vector. Let $\mathbf{n}_0$ and $\mathbf{n}_1$ be the normal vectors to the initial and final planes. Let R_0 be the (smallest) rotation in the $\mathbf{e}_3$–$\mathbf{n}_0$ plane which takes $\mathbf{n}_0$ to $\mathbf{e}_3$ and R_1 the analogous rotation in the $\mathbf{e}_3 - \mathbf{n}_1$ plane which takes $\mathbf{e}_3$ to $\mathbf{n}_1$. (If the normal vector $\mathbf{n}_i$ is coincident with $\mathbf{e}_3$ then its rotation R_i is the identity.) Since R takes $\mathbf{n}_0$ to $\mathbf{n}_1$ it can be written in the form

$$R = R_1 R_{\mathbf{J}_0} R_0 \tag{1}$$

† email: rmont@cats.ucsc.edu

where $R_{\mathbf{J}_0}$ is a rotation about the $\mathbf{J}_0 = \mathbf{e}_3$ axis by some angle $\Delta\theta$. Our main result is the following integral formula for this angle

$$\Delta\theta = \int_0^{t_1} \omega(t)\mathrm{d}t + \int\int_D \Omega. \tag{2}$$

We will take the remainder of this subsection to describe the integrands of this formula. The formula itself is another example of a 'reconstruction formula' ([6, 5, 10]) and as such is closely related to Berry phase formulas [16]. We will expand on this at the end of this subsection.

The first integral, $\int \omega \mathrm{d}t$ is called a 'dynamic phase' in the Berry phase literature. The integrand ω represents the instantaneous angular velocity of the moving triangle $q(t)$ about the axis $\mathbf{J}_0$. It is given by

$$\omega(t) = \mathbf{e}_3 \cdot \mathbb{I}(q(t))^{-1}\mathbf{J}_0 \tag{3}$$

which also equals $J_0\mathbf{e}_3 \cdot \mathbb{I}(q(t))^{-1}\mathbf{e}_3$. Here, $\mathbb{I}(q)$ is the instantaneous moment of inertia tensor of the weighted triangle q (see (7) below) and recall that $\mathbf{J}_0 = J_0\mathbf{e}_3$. Physically speaking, $\mathbb{I}(q(t))^{-1}\mathbf{J}_0$ is the angular velocity which a weighted triangle would have, if its angular momentum were $\mathbf{J}_0$, and if it were to be frozen at the shape $q(t)$ which it had at time t. The time t_1 of integration is the duration of the motion.

The second integral is of the type called a 'geometric phase' in the Berry phase literature. It is the integral of a two-form, Ω, over a disc, D, in a certain 'reduced space' which we call Z. The two-form, Ω, is closed and is independent of the choice of potential defining the three-body dynamics. It lives on a four-dimensional 'reduced configuration space', denoted $Z = Z(\mathbf{J}_0)$ which we will now describe. The configuration space for the three-body problem in three-space, with centre of mass fixed, is a six-dimensional Euclidean space denoted by Q. Its elements will be thought of as triangles, with vertices $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ having masses m_1, m_2, m_3 and with centre of mass at the origin. The group, $G = SO(3)$ of rotations acts on Q by isometries. The group $\mathbb{R}^+$ of dilations and the subgroup $SO(2) \subset G$ of rotations about the $\mathbf{J}_0$-axis also act on Q. It will be convenient, in fact essential, for us to enlarge Q to form $\tilde{Q}$, the space of oriented triangles. An element of $\tilde{Q}$ is a triangle $q \in Q$, together with a unit vector $\mathbf{n}$ which is orthogonal to the subspace spanned by the vertices of q. Let 0 denote the triple collision $\mathbf{q}_1 = \mathbf{q}_2 = \mathbf{q}_3 = 0$. Then Z is, by definition, the quotient space

$$Z = (\tilde{Q} \setminus 0)/(SO(2) \times \mathbb{R}^+).$$

The shape space or sphere, discussed earlier, is the quotient

$$S = (\tilde{Q} \setminus 0)/(G \times \mathbb{R}^+) = \text{shape sphere}.$$

The inclusion $SO(2) \subset G = SO(3)$, naturally induces a projection $\pi : Z \to S$ with typical fibre $\pi^{-1}(s)$ equal to $G/SO(2)$, which is a two-sphere. Thus Z is a two-sphere bundle over the two-sphere S. We urge the reader to look at the first figure which is a picture of Z and various of its features. Theorem 1 below asserts that Z is the nontrivial two-sphere bundle over the two-sphere, S. (There are exactly two S^2 bundles over S^2. One is the trivial bundle. Z is the other one.)

The fibring spheres of Z are interpreted as instantaneous versions of the body angular momentum sphere which occurs in the description of the motion of a free rigid body. 'Instantaneous' refers to the instantaneous shape of the triangle. A point (x, y, z) on a fibring sphere can be thought of as the fixed total angular momentum vector $\mathbf{J}_0$ viewed with respect to a coordinate system attached to the moving triangle. Let $\{\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3\}$ be the moving orthonormal frame defining the moving coordinate axes. Then $(x, y, z) = (\mathbf{J}_0 \cdot \mathbf{U}_1, \mathbf{J}_0 \cdot \mathbf{U}_2, \mathbf{J}_0 \cdot \mathbf{U}_3)$. We will always take the third vector in our frame to be the

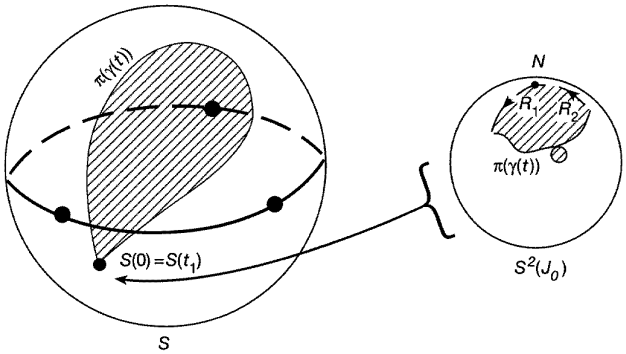

Figure 1. The reduced configuration space S_{J_0} with the projected curve of integration $\pi o \gamma(t)$. The space is locally the product of the shape sphere S and a momentum sphere $S^2(J_0)$. The collinear configurations (equator) and double collisions are indicated on S by a thick curve and the dots on this curve

normal $\mathbf{n}$ to the triangle: $\mathbf{U}_3 = \mathbf{n}$. A good choice of such a moving frame is the principal axis frame. This is a right-handed orthonormal frame which diagonalizes the instantaneous inertia tensor $\mathbb{I}$ of the triangle. (The normal vector $\mathbf{n} = \mathbf{U}_3$ corresponds to the eigenvalue I of the inertial tensor. For a typical configuration the eigenvalues of $\mathbb{I}$ are distinct, and there are actually three such frames, differing from each other by signs.)

The two-form, Ω, is defined as follows. Begin with the natural mechanical connection $\mathbf{A}$ ([1, 6, 9]). The geometry behind this form is reviewed in subsection 1.3 and subsection 2.6. It can be defined by the equation $\mathbf{A} = \mathbb{I}^{-1}\mathbf{J}$ where the angular momentum, $\mathbf{J}$, is thought of as a one-form on configuration space with values in $\mathbb{R}^3$. $\mathbf{A}$ is defined everywhere except where $\mathbb{I}$ is not invertible, which is precisely at the collinear configurations (see subsections 1.4 and 1.5). After pulling $\mathbf{A}$ back to $\tilde{Q}$ it is smooth everywhere. Now consider its component α along the $\mathbf{e}_3$ axis, and form the exterior derivative $\mathrm{d}\alpha$. This yields a two-form on $\tilde{Q}$. Next, observe that this two-form is is basic with respect to the action of $\mathbb{R}^+$ and $SO(2)$, which is to say, it is the pull-back of a two-form on the quotient space Z. This two-form on Z is the two-form Ω.

To give a formula for the two-form Ω on Z we use standard spherical coordinates (ϕ, θ) on spheres $S^2(R)$, as well as coordinates (z, θ) where

$$z = \cos(\phi)$$

is the normalized height of a point above the equatorial circle $\phi = \pi/2$ so that Rz is the usual height, and $-1 \leqslant z \leqslant 1$. These spherical coordinates, together with the above local trivialization of Z into the product of two S^2s induces coordinates $(z_1, \theta_1, z_2, \theta_2)$ on Z. We will show that

$$\Omega = -\{\tfrac{1}{2}\mathrm{d}(z_1 z_2) \wedge \mathrm{d}\theta_1 + \mathrm{d}z_2 \wedge \mathrm{d}\theta_2\}. \tag{4}$$

The height coordinate z_1 on the shape sphere is proportional to the area A of the triangle. (See the appendix for a derivation.) Explicitly,

$$z_1 = 4\sqrt{\frac{m_1 m_2 m_3}{m_1 + m_2 + m_3}}\frac{A}{I}.$$

Here

$$I = m_1\|\mathbf{q}_1\|^2 + m_2\|\mathbf{q}_2\|^2 + m_3\|\mathbf{q}_3\|^2 = \|q\|^2$$

is its polar moment of inertia, and also the square of the Euclidean norm on Q. And $A = \frac{1}{2}\mathbf{n} \cdot (\mathbf{q}_2 - \mathbf{q}_1) \times (\mathbf{q}_3 - \mathbf{q}_1)$ is the oriented area of the triangle.

The height coordinate on the fibring spheres is given by

$$z_2 = \frac{1}{J_0}\mathbf{J}_0 \cdot \mathbf{n}.$$

It is the component of the total angular momentum normal to the triangle.

The angular coordinates θ_1, θ_2 do not have such simple meanings, and are more arbitrary. In particular, the coordinate θ_2, and the local splitting of Z into the product of spheres depends on the choice of local frame $\{\mathbf{U}_i\}$ (choice of gauge) for the moving triangle $q(t)$. The choice which yields our coordinate formula for Ω is the principal axis frame discussed above.

A three-body motion without triple collision has natural projections to a curve in Z and to a curve in S. Our assumption that the initial and final triangles are similar means that either this path in S is closed, or that its endpoints are related by reflection about the equator, E. The latter situation occurs when the rotation which takes one triangle to (a homothety of) the other, takes the initial normal vector to the negative of the final normal vector. For simplicity, we will suppose that we are in the first case, where the path in S is closed. The path in Z need not be closed, but there is a canonical way to close it. This is depicted in the first figure as the arcs on the fibring spheres labelled by R_1 and R_2. The disc, D, over which we integrate the two-form, Ω, is any disc in Z bounding the resulting closed curve. (Z is simply connected.)

To summarize, the data needed for our formula are

- the total angular momentum $\mathbf{J}_0$,
- the initial and final normals $\mathbf{n}_0$ and $\mathbf{n}_1$ to the triangle planes,
- the angular speed $\omega(t)$ of (3),
- the reduced curve in Z.

The last two pieces of data can be recovered from a curve in the space

$$Q_{\mathbf{J}_0} = (\tilde{Q} \setminus 0)/SO(2) = \text{ reduced configuration space.}$$

This space fibres over Z with fibre coordinatized by I and describing the overall size of the triangle. This size is needed to define $\omega(t)$. Newton's equations for fixed angular momentum $\mathbf{J}_0$ can be written as a mixed first and second order differential equation on this reduced configuration space [11, 13]. The space itself is closely related to the usual symplectic reduced space. Our formula, then, provides a means of recapturing the original dynamics on Q if the reduced dynamics is known.

1.2. The planar case

In the planar case our question is simpler and has been solved several times before [4, 1, 9, 3]. A single angle describes the rotation relating two similar planar triangles. This angle $\Delta\theta$ is described by the same formula as above, which simplifies as follows. The integrand, ω, for the dynamic phase becomes $\frac{1}{I(t)}J_0$. To obtain the planar two-form, simply set $z_2 = 1$ in the formula for Ω above. This constraint describes the embedding of the planar problem in the spatial one: the triangle's normal is aligned with the angular momentum vector. (And, if it is so initially, it is for all time.) Now $\mathrm{d}z_2 = 0$ and the two-form integrand becomes $\Omega = -\frac{1}{2}\mathrm{d}z_1 \wedge \mathrm{d}\theta_1 = +\frac{1}{2}\sin(\phi_1)\mathrm{d}\phi_1 \wedge \mathrm{d}\theta$ on the shape sphere, which is twice the area form on the sphere S of radius $1/2$. It is the unique spherically symmetric two-form on the sphere whose total integral is 2π.

1.3. The structure of the calculation and some history

Our derivation of the reconstruction formula is quite similar to our earlier derivation [10] of a reconstruction formula for rigid body motion. We construct a closed loop γ and a one-form α in the 'regularized' three-body configuration space $\tilde{Q}$ such that when we apply Stoke's theorem to the line integral $\int_\gamma \alpha$ and evaluate various terms, we obtain our formula. We construct the loop, γ, by concatenating the three-body motion $q(t)$ defined by Newton's equations with several other arcs obtained by rotating or scaling. The one-form, α, was described in subsection 1.1. It is the $\mathbf{e}_3$ component of the 'natural mechanical connection' $\mathbf{A}$, pulled back to $\tilde{Q}$.

The connection $\mathbf{A}$ first appears in Guichardet [1]. It has been argued that it was discovered by Smale, who had a formula for something like our one-form $\alpha_{\mathbf{J}_0}$ [17]. The connection form was later used by Iwai [4] and rediscovered by Shapere and Wilczek [16]. I used it in [9] in studying the falling cat problem. The decomposition of velocities which $\mathbf{A}$ defines is known to modern celestial mechanists as 'Saari's decomposition'. See [14], and also [15]. See also the survey of Reinsch and Littlejohn [13] for explicit descriptions and formulae and some applications to quantum mechanics.

The calculation would be conceptually simpler if we could work directly on Q. Unfortunately, the connection one-form becomes singular, and is undefined, at the collinear configurations. A collinear configuration, also called an eclipse or syzygy is by definition a three-body configuration in which the three masses lie on a single line. Hsiang's device [3] of introducing the space of $\tilde{Q}$ of oriented triangles allows us to get rid of this singularity. It appears to be necessary to work on $\tilde{Q}$ to obtain our formula.

The essential fact that the space S of similarity classes of triangles is the two-sphere is well-known to modern celestial mechanics. See for example Moeckel's beautiful survey [8].

1.4. Collinear configurations and Hsiang's regularization

The configurations with extra rotational symmetry are precisely the collinear configurations $C \subset Q$. At any noncollinear configuration q the rotation group, G acts freely, meaning that $gq = q$ implies $g = 1$. Thus $Q \to Q/G$ fails to be a principal bundle exactly at the collinear configurations, and hence the connection form $\mathbf{A}$, well-defined away from C, must become singular at C. (The normalization condition, satisfied by any connection must fail at points with nontrivial continuous symmetries.)

Hsiang's desingularization $\beta : \tilde{Q} \to Q$ allows us to extend the connection one-form and hence its component α to the collinear configurations. More precisely, the form $\beta^*\mathbf{A}$ extends smoothly to $\beta^{-1}(C) \setminus \{0\}$, although $\mathbf{A}$ does not extend to $C \subset Q$. By abuse of notation we will refer to the pull-backs $\beta^*\mathbf{A}$, $\beta^*\alpha$ and $\beta^*\mathrm{d}\alpha$ simply as $\mathbf{A}, \alpha$, and $\mathrm{d}\alpha$. As stated in subsection 1.1, the form $\mathrm{d}\alpha$ on $\tilde{Q}$ is the pull-back of the form Ω on Z which is the form of our reconstruction formula (2). The desingularization $\tilde{Q}$ has the added advantage of making various quotient spaces, such as $\tilde{Q}/G$, nonsingular smooth manifold without boundary away from the triple collision point 0.

At first glance it may appear to the reader that any solution curve with a syzygy must have angular momentum $\mathbf{J} = 0$, and that consequently we can ignore syzygies (and thus the desingularization) for the nonplanar problem. However, this is not the case. Physically, one can see this in the lunar problem where eclipses are possible even when the moon–earth orbital plane and sun–moon orbital plane do not coincide. Mathematically we can see this possibility as follows. Suppose that a syzygy occurs at the configuration $q_0 = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$.

Then we may write: $\mathbf{q}_i = s_i\mathbf{e}$ where $\mathbf{e}$ is a fixed unit vector and the s_i are scalars with

$$\Sigma m_i s_i = 0. \tag{5}$$

Now write $\dot{q}_0 = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for the velocities at syzygy. We have

$$\Sigma m_i \mathbf{v}_i = 0 \tag{6}$$

since we are assuming that we are in a centre of mass frame. The angular momentum is

$$\mathbf{J} = \Sigma m_i \mathbf{q}_i \times \mathbf{v}_i = \mathbf{e} \times (\Sigma m_i s_i \mathbf{v}_i).$$

Most solutions (s_i), $(\mathbf{v}_i)$ to the above constraint equations will satisfy $\mathbf{J} \neq 0$.

It follows immediately from this formula for $\mathbf{J}$ that it is impossible for the axis $\mathbf{e}$ of a collinear configuration to be equal to $\mathbf{e}_3$, the axis of the angular momentum vector $\mathbf{J} = \mathbf{J}_0$. The only singular points of $Q/SO(2)$ occur precisely at collinear configurations aligned with $\mathbf{e}_3$ and we have just seen that these are not physically realizable. This suggests that it may be possible to carry out all of our calculations directly on $Q/SO(2)$, thus getting rid of the device $\tilde{Q}$ of oriented triangles. Question: can our reconstruction formula and calculations be reformulated solely in terms of $Q/SO(2)$, without relying on oriented triangles? If the answer is 'yes' the calculation and formula will have to be different, since the form α does not extend to $Q/SO(2)$, even after we have deleted these 'aligned' collinear configurations.

Although syzygies can occur in the three-body problem they are rare. The set C of all collinear configurations has codimension 2 within the configuration space Q for the spatial problem, and so we expect that 'most' solution curves will miss C. The projections of these typical nonecliptic curves to the shape sphere S will lie entirely in one (open) hemisphere defined by the equator $E = \pi(C)$. Even when a syzygy does occur along a solution path, if its projected path to S is closed then the number of syzygies must be even since it must cross the equator E an even number of times. Consequently we expect such paths to have codimension 4 within the space of all solutions. On the other hand, the collinear configurations for the planar problem has codimension 1 and we expect that 'most' (an open set's worth) of solution curves admit syzygies.

The shape sphere S is the same for both the planar and spatial problem, thanks to Hsiang's desingularization. In both cases the projection $E = \pi(C)$ of the syzygies forms the equator of S, a set of codimension 1. The apparent discrepancy between the codimensions of C and $\pi(C)$ in the spatial case is a result of Hsiang's desingularization: the fibre $\pi^{-1}(s)$ has dimension 3 if $s \notin E$ (the fibre is a copy of the rotation group) and it has dimension 2 if $s \in E$ (the fibre is a two-sphere).

1.5. Outline of the paper

In the next section we introduce some notation and constructions basic to our goals. Then we present the basic theorems regarding the metric and topological structure of the quotients S, Z and some intermediate quotients. We also describe more carefully the two-form Ω. In subsection 3 we prove these theorems. The proofs are based on restricting to the planar three-body problem in which case the Hopf fibration arises naturally. In the final section, we prove our reconstruction formula in the manner outlined above. In appendix A we derive the fact that the shape sphere S is a sphere of radius $1/2$. In appendix B we provide an alternative calculation of the basic two-form Ω, one which shows that the principal axis frame is the gauge in which our formulae are valid.

2. Constructions, notation and theorems

2.1. Basic notation

The three-body configuration space Q consists of the set of all triples of vectors $q = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$, $\mathbf{q}_a \in \mathbb{R}^3$ whose centre of mass is at the origin: $\Sigma m_a \mathbf{q}_a = 0$. The positive numbers m_a are the particle masses. We prefer to view Q as the space of weighted triangles in space. In any case it is is a six-dimensional Euclidean vector space with squared norm $I(q) := \|q\|^2 := m_1\|\mathbf{q}_1\|^2 + m_2\|\mathbf{q}_2\|^2 + m_3\|\mathbf{q}_3\|^2$, also called the polar moment of inertia. The instantaneous kinetic energy T (not to be confused with the 'T' of tangent bundle!) of a path $q(t)$ is defined to be $T := \frac{1}{2}\|\dot{q}\|^2$. where

$$\dot{q} = (\dot{\mathbf{q}}_1, \dot{\mathbf{q}}_2, \dot{\mathbf{q}}_3)$$

denotes the time derivative of the path.

The instantaneous total energy of a motion is $E = T + V$, where $V = -\Sigma_{i<j} \frac{m_i m_j}{\|\mathbf{q}_i - \mathbf{q}_j\|}$ is the usual Newtonian gravitational potential energy. The choice of potential actually plays no role in our analysis. All of our results hold, with the same proofs, for any potential which is rotationally invariant. Thus any potential which is a function of the interparticle distances alone will work. A three-body motion is a solution $q(t)$ to Newton's equations: $m_a \frac{d^2}{dt^2}\mathbf{q}_a = -\boldsymbol{\nabla}_a V$, $a = 1, 2, 3$. The total energy E and the total angular momentum $\mathbf{J} = \Sigma m_a \mathbf{q}_a \times \dot{\mathbf{q}}_a$ are constant along any three-body motion.

The moment of inertial tensor $\mathbb{I}(q)$ of a weighted triangle q is the symmetric non-negative 3×3 matrix defined by

$$\omega \cdot \mathbb{I}(q)\omega = \|\omega \times q\|^2 \tag{7}$$

where $\omega \times q = (\omega \times \mathbf{q}_1, \omega \times \mathbf{q}_2, \omega \times \mathbf{q}_3)$ denotes the infinitesimal rotation of the triangle q with angular velocity ω. $\mathbb{I}$ encodes that part of the metric on Q in the direction of the G-orbits. It satisfies the equivariance relations: $\mathbb{I}(\lambda R q) = \lambda^2 R \mathbb{I}(q) R^T$ where $\lambda \in \mathbb{R}^+$ is a homothety, or dilation and $R \in G = SO(3)$ is a rotation. (The inertia tensor can be expressed by the formula $\mathbb{I}(q) = I(Id.) - \mathbb{M}$ where $\mathbb{M} = \Sigma m_a \mathbf{q}_a \otimes \mathbf{q}_a$ is the standard inertia tensor, $Id.$ is the identity matrix, and $I = tr(\mathbb{M})$ is the polar moment of inertia.)

2.2. Oriented triangles

The idea in this section is due to Hsiang [3]. We will need to choose a normal vector $\mathbf{n}(t)$ to the plane of our moving triangle $q(t)$.

Definition 1. *An oriented triangle is a pair $\tilde{q} = (q, \mathbf{n})$ with $q \in Q$ and $\mathbf{n} \in \mathbb{R}^3$ a unit vector normal to the subspace of $\mathbb{R}^3$ spanned by the vertices $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ of the triangle q. (Since the centre of mass is 0 these three position vectors are always linearly dependent.) The set of oriented triangles will be denoted by $\tilde{Q}$.*

If the vertices of our triangle q span a plane then it has two possible orientations $\tilde{q} = (q, \pm\mathbf{n})$ where $\pm\mathbf{n}$ are either of the two normals to this plane. If the triangle lies in a single line an orientation for it is any vector $\mathbf{n}$ on the unit circle in the plane orthogonal to this line. Such triangles are called collinear configurations. If $q = 0$ is the triple collision point then $\mathbf{n}$ is any point on the two-sphere.

Lemma 1. *$\tilde{Q}$ is a smooth algebraic variety. Away from the triple collision, the natural projection $\tilde{Q} \to Q$ is a branched cover, branched over the collinear configurations. The rotation group $G = SO(3)$ acts freely on $\tilde{Q}$ away from the triple collision point.*

Proof. $\tilde{Q}$ is the algebraic subvariety of $Q \times S^2$ defined by the two equations $\mathbf{q}_1 \cdot \mathbf{n} = 0$, $\mathbf{q}_2 \cdot \mathbf{n} = 0$. The differential of these defining functions is full-rank everywhere. Apply the implicit function theorem. The other statements are obvious.

It follows from this lemma that any smooth function or covariant tensor on Q can be lifted to $\tilde{Q}$. Examples are $V, \mathbb{I}$ and the Riemannian metric. The lift will be denoted by the same symbol as the original. The lifted metric fails to be positive definite along the branching locus. The three-body equations themselves also lift to $\tilde{Q}$:

Lemma 2. *Any three-body motion $q(t)$ which does not consist entirely of collinear configurations has a unique oriented lift $\tilde{q}(t) \in \tilde{Q}$ passing through a given initial noncollinear oriented triangle.*

The proof is obvious.

2.3. The reduced configuration space, the shape sphere, and other quotient spaces

Let

$$G(\mathbf{J}_0) = SO(2) \subset G$$

denote the one-parameter subgroup of rotations about the angular momentum axis $\mathbf{J}_0$. The quotient space $Q/G(\mathbf{J}_0)$ is singular, even away from the triple collision, due to the presence of extra symmetry at collinear configurations. The introduction of the space $\tilde{Q}$ of oriented triangles regularizes this quotient away from the triple collision.

Definition. *The reduced configuration space is the quotient space*

$$Q_{\mathbf{J}_0} = \tilde{Q}/G(\mathbf{J}_0),$$

with corresponding projection denoted by

$$\pi_{\mathbf{J}_0} : \tilde{Q} \to Q_{\mathbf{J}_0}.$$

The reduced motion corresponding to the oriented three-body motion $\tilde{q}(t)$ is the projection $\pi_{\mathbf{J}_0}(\tilde{q}(t))$ of this curve to the reduced configuration space.

The reduced configuration space is essentially a cone over the space Z which plays a central role in our reconstruction formula. In order to show this and in order to get a good picture of both spaces, we will also need to understand various other quotient spaces. Set

$$\bar{Q} := \tilde{Q}/G = \text{congruence classes of oriented triangles},$$

$$Q/G = \text{congruence classes of triangles}.$$

The action of $\lambda \in \mathbb{R}^+$ scales each triangle by the factor λ and scales distances on Q by this same factor. The polar moment (squared norm) I is a G-invariant function which is homogeneous of degree 2. Let

$$S^5 = \{I = 1\} \subset Q$$

denote the five-sphere in the Euclidean space Q and

$$\tilde{S}^5 \subset \tilde{Q}$$

be the corresponding preimage of this sphere under the branched cover $\tilde{Q} \to Q$. Also let

$$\tilde{Q}^* = \tilde{Q} \setminus \{0\},$$

and

$$Q^* = Q \setminus \{0\}.$$

(The 0 in $\tilde{Q}$ represents the two-sphere of oriented triple collisions.) Then we have natural identifications:

$$Q^*/\mathbb{R}^+ \cong S^5$$

and

$$\tilde{Q}^*/\mathbb{R}^+ \cong \tilde{S}^5 \subset \tilde{Q}.$$

The space $Q^*/(G \times \mathbb{R}^+)$ of similarity classes of triangles is naturally isomorphic to

$$S_+ := S^5/G \subset Q/G$$

and the space $\tilde{Q}^*/(G \times \mathbb{R}^+)$ of similarity classes of oriented triangles is isomorphic to

$$S := \tilde{S}^5/G \subset \tilde{Q}/G.$$

Define

$$Z = \tilde{S}^5/G(\mathbf{J}_0) \subset Q_{\mathbf{J}_0}.$$

Corresponding to these spaces we have various projections, $Q^* \to S_+$, $\tilde{Q}^* \to S$, $\tilde{S}^5 \to Z$, et cetera, denoted by π or $\pi_{\mathbf{J}_0}$. Note that the fibres of

$$Z \to S$$

are the two-spheres:

$$\pi_{\mathbf{J}_0}^{-1}(pt.) = S^2(J_0) = G/G(\mathbf{J}_0).$$

Theorem 1. *S is a two-sphere which we call the shape sphere. The projection $\tilde{S}^5 \to S$ is the nontrivial principal $SO(3)$ bundle over the two-sphere. The projection $Z \to S$ is the associated nontrivial two-sphere bundle over the two-sphere. The reduced configuration space, $Q_{\mathbf{J}_0}$, minus the triple collision, is diffeomorphic to $Z \times (0, \infty)$ where the second factor is parameterized by I.*

Explanation. If G is a connected Lie group, then the equivalence classes of principal G-bundles over an n-sphere are parameterized by the homotopy group $\pi_{n-1}(G)$. In our case this homotopy group is $\pi_1(SO(3))$ which is the two-element group. The nontrivial G-bundle over S^2 can be realized as follows. Identify the two-sphere with the complex projective line CP^1. Let $\gamma \to CP^1$ denote the canonical complex line-bundle and $\epsilon = CP^1 \times \mathbb{R}$ the trivial real line bundle. Form the rank 3 real vector bundle $E = \gamma \oplus \epsilon$. This is an oriented vector bundle with a natural fibre-inner product. Then the nontrivial bundle, our $\tilde{S}^5$, is the bundle of oriented orthonormal frames for E. And the nontrivial sphere bundle, our Z, is the unit sphere bundle of E.

2.4. Metric nature of the quotients

Q, being a Euclidean space, is a metric space. G acts on it by isometries, so that the quotient space Q/G of congruence classes of weighted, centred triangles, inherits a metric. This quotient metric—sometimes called the orbital distance metric—is defined by declaring that the distance between two points in the quotient is the distance between the corresponding orbits in the original space. The dilations $\mathbb{R}^+$ act on Q and commute with the G action so they induce an action on the quotient as well.

The cone over a Riemannian manifold $(X, \mathrm{d}s^2)$, possibly with boundary, is the topological space $(X \times [0, \infty))/(X \times \{0\})$ with associated Riemannian metric $\mathrm{d}\lambda^2 + \lambda^2 \mathrm{d}s^2$. Here λ is the real parameter in $[0, \infty)$. The quotient by '$X \times \{0\}$' means that we crush (identify) $X \times \{0\}$ to a single point, called the 'cone point'. The metric tensor and manifold structure becomes singular there.

Theorem 2. *The metric space Q/G of congruence classes of triangles is a cone over the space S_+ of similarity classes. The cone point corresponds to the triple collision. S_+ is isometric to the closed upper hemisphere of radius one-half. The equator represents collinear configurations. The dilation parameter is*

$$\lambda = \sqrt{I}.$$

Replacing Q by $\tilde{Q}$ resolves the singularity corresponding to the collinear configurations. The pull-back to $\tilde{Q}$ of the metric on Q fails to be a metric over the collinear configurations: it takes no energy to rotate a line segment about its axis. However, dividing by the G-action kills these null directions so we again get a metric on the quotient $\tilde{Q}/G$.

Theorem 3. *The metric space $\tilde{Q}/G$ of congruence classes of oriented triangles is a cone over the space S of similarity classes of oriented triangles. The cone point corresponds to the triple collision. S is isometric to to the two-sphere of radius one-half. The equator corresponds to the collinear configurations. The height coordinate above the equator is*

$$\frac{1}{2} z_1 = 2\sqrt{\frac{m_1 m_2 m_3}{m_1 + m_2 + m_3}} \frac{\Delta}{I}.$$

where Δ is the oriented area of the triangle:

$$\Delta = \frac{1}{2} \mathbf{n} \cdot (\mathbf{q}_2 - \mathbf{q}_1) \times (\mathbf{q}_3 - \mathbf{q}_1).$$

The map $S \to S_+$ of the sphere to the hemisphere which is induced by the branched cover $\tilde{Q} \to Q$ corresponds to the quotient map obtained when we identify the hemisphere with the quotient space obtained by identifying points of the sphere related by the reflection about this equator.

2.5. Relation with the symplectic reduced space

This section is included to connect the constructions of the previous two sections with the symplectic reduced space construction. The results here are not used in arriving at our formula, but may shed some light on it.

Consider the general situation of a compact Lie group G acting freely on a manifold Q, and so also on T^*Q. Recall that the symplectic reduced space at the point $\mathbf{J}_0 \in \mathrm{Lie}(G)^*$ is the sub-quotient space $J^{-1}(\mathbf{J}_0)/G(\mathbf{J}_0)$ where $J : T^*Q \to \mathrm{Lie}(G)^*$ is the momentum map of the action, $\mathbf{J}_0$ is a particular fixed element of $\mathrm{Lie}(G)^*$, and $G(\mathbf{J}_0) \subset G$ is its isotropy group (relative to the dual of the adjoint action). This symplectic reduced space is diffeomorphic to the fibre product: $T^*(Q/G) \times_f (Q/G(\mathbf{J}_0))$ over the quotient Q/G. This follows directly from [12], or [18], together with the fact that $Q/G(\mathbf{J}_0)$ is naturally identified with the co-adjoint orbit bundle $Q \times_G (G/G(\mathbf{J}_0)) \subset Q \times_G \mathrm{Lie}(G)^*$ over Q/G. (See also the chapter in [2] on minimal coupling.) In a case such as the three-body problem where the underlying dynamics can be described by a second order equation on Q, the T^*-part of a reduced solution curve in the reduced space can be recovered from the derivative of the projection

of that curve to Q/G. It follows that the entire reduced curve can be recovered from its projection to $Q/G(\mathbf{J}_0)$. Thus it makes sense to call $Q/G(\mathbf{J}_0)$ the reduced configuration space at $\mathbf{J}_0$.

2.6. The connection

Our quotients inherit various tensorial objects besides metrics. In order to describe them we proceed generally. Suppose again that we are given a Riemannian manifold Q and a group G of isometries of Q acting freely. From this data we can form:

- a metric on the quotient
- a connection for the principal G-bundle $Q \to Q/G$
- a fibre inner-product on the adjoint bundle $Q \times_G \mathrm{Lie}(G) \to Q/G$ of Lie algebras over the quotient.

The metric on the quotient (orbital distance metric) we have described.

To define the connection, we define its horizontal space.

Definition. *The horizontal space at $q \in Q$ is the orthogonal complement at q to the group orbit through q. The associated connection form $\mathbf{A} : TQ \to \mathrm{Lie}(G)$ is called the natural connection.*

The metric on Q/G is a Riemannian one. Its metric tensor is obtained by identifying the tangent space at $\pi(q)$ with the horizontal space at q. With this definition, the projection $Q \to Q/G$ has the structure of a Riemannian submersion.

To define the fibre inner-product on the adjoint bundle, let

$$\sigma(q) : \mathrm{Lie}(G) \to T_q Q$$

denote the infinitesimal generator of the group action:

$$\sigma(q)(\omega) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} \exp(\epsilon\omega) q|_{\epsilon=0}.$$

Then set

$$\|\omega\|_q^2 = \|\sigma(q)\omega\|_Q^2.$$

This defines the fibre inner product. Fix a bi-invariant inner product $\cdot$ on $\mathrm{Lie}(G)$. Using the inner products we can construct the transpose

$$\sigma^T(q) : T_q Q \to \mathrm{Lie}(G).$$

The map $(q, \dot{q}) \to \sigma^T(q)(\dot{q})$ is the Noether conserved quantity, or, after we identify $\mathrm{Lie}(G)$ and TQ with their duals using the inner products, it is the momentum map $J : T^*Q \to \mathrm{Lie}(G)^*$. The fibre-inner product can also be written

$$\omega \cdot \mathbb{I}(q)\omega = \|\omega\|_q^2,$$

thus defining the moment of inertia tensor $\mathbb{I}$. We have $\mathbb{I}(q) = \sigma^T(q)\sigma(q)$. We now have the universal formula for the connection form associated to this situation:

$$\mathbf{A}(q) = \mathbb{I}(q)^{-1} \circ \sigma^T(q).$$

In our situation $\sigma^T = \mathbf{J}$ is the angular momentum, viewed as a one-form with values in $\mathbb{R}^3$ (the Lie algebra of $G = SO(3)$):

$$\mathbf{J} = \Sigma m_a \mathbf{q}_a \times d\mathbf{q}_a.$$

$\mathbb{I}$ is of course our moment of inertia tensor $\mathbb{I}$. The connection form is then given by

$$\mathbf{A}(q) = \mathbb{I}(q)^{-1} \circ \mathbf{J}(q).$$

The horizontal space is the space of infinitesimal deformations with zero angular momentum. This gives us a physical picture of what it means for a curve to be horizontal, and of the length of a path in one of the quotient spaces S, Q/G, et cetera.

$\mathbf{A}$ has a nice physical interpretation. If $q(t)$ is a three-body motion then $\mathbf{A}(q(t)(\dot{q})$ is the 'best' choice of assignment of an angular velocity ω to the motion, given the fact that this motion need not be a rigid motion. If it does happen to be a rigid motion, with infinitesimal angular velocity ω, then $\mathbf{A}(q(t)(\dot{q}) = \omega$.

All of this falls apart at the collinear configurations. Rotation about the axis, $\mathbf{e}$ of such a configuration q leaves the configuration fixed. It follows that the infinitesimal generator $\sigma(q)(\mathbf{e}) = 0$. Consequently the normalization condition, $\mathbf{A}(q)\sigma(q)(\mathbf{e}) = \mathbf{e}$ fails at the collinear configurations. The connection, $\mathbf{A}$ cannot be extended to the set C of collinear configurations. This is the main reasons for introducing $\tilde{Q}$. The action of the rotation group on $\tilde{Q}$ is free away from the triple collision, and consequently the connection form, $\mathbf{A}$ is well-defined everywhere on $\tilde{Q}$. More precisely, the pull-back by $\beta : \tilde{Q} \to Q$ of the mechanical connection, $\mathbf{A}$ extends smoothly to the 'branching locus' $\beta^{-1}(C)$. By slight abuse of notation, we also refer to this form as $\mathbf{A}$.

Definition. *The form α is the component of $\mathbf{A}$ on $\tilde{Q}$ along the axis $\mathbf{e}_3$ of our fixed angular momentum vector $\mathbf{J}_0$. That is:*

$$\alpha = \mathbf{e}_3 \cdot \mathbf{A}.$$

It is a one-form on on $\tilde{Q}^* = \tilde{Q} \setminus 0$.

Remark 1. *The pull-back of α along an oriented three-body motion $\tilde{q}(t)$ satisfies*

$$\tilde{q}^*\alpha = \omega(t)\mathrm{d}t$$

where $\omega(t) = \omega(q(t))$ is the instantaneous angular velocity of that motion about the $\mathbf{J}_0$-axis, that is to say, the first integrand of our main formula (2).

Remark 2. *Away from the triple collision, the natural projection $\pi_{\mathbf{J}_0} : \tilde{Q} \to Q_{\mathbf{J}_0}$ has the structure of a principal circle bundle, the circle being $G(\mathbf{J}_0)$. Its associated connection one-form is α.*

Theorem 4. *The form $\mathrm{d}\alpha$ pushes down to a two-form Ω on Z. This is the form Ω described in the introduction and given by the explicit formula (4) above.*

3. Planar configurations and proofs

3.1. The Hopf fibration in planar configurations

A planar configuration is a triangle lying in the plane perpindicular to the angular momentum vector $\mathbf{J}_0$. If the triangle is oriented we will take its normal to be parallel to $\mathbf{J}_0$: $\mathbf{J}_0 \cdot \mathbf{n} = J_0 > 0$. The set Q_{planar} of planar configurations forms a four-dimensional Euclidean subspace of the full configuration space. The action of the circle group $G(\mathbf{J}_0)$ on Q_{planar} is isomorphic to the action of the circle on $\mathbf{C}^2$ which takes (ζ_1, ζ_2) to $(e^{i\theta}\zeta_1, e^{i\theta}\zeta_2)$.

The intersection of the five-sphere $\{I = 1\}$ with Q_{planar} forms a round three-sphere, denoted either $\tilde{\Sigma} \subset \tilde{Q}$ or $\Sigma \subset Q$. These three-spheres are diffeomorphic under $\beta : \tilde{Q} \to Q$

due to the unique choice of orientation. Their quotients by $G(\mathbf{J}_0)$ are isometric to the two-sphere of radius $\frac{1}{2}$, which is the shape sphere S. Thus:

$$G(\mathbf{J}_0) \to \tilde{\Sigma} \to \Sigma/G(\mathbf{J}_0) = S$$

and

$$G(\mathbf{J}_0) \to \Sigma \to \Sigma/G(\mathbf{J}_0) = S$$

are isometric as Riemannian submersions to the standard Hopf fibration:

$$S^1 \to S^3(1) \to S^2\left(\tfrac{1}{2}\right).$$

3.2. Proof of theorems

Proof of theorem 1. Consider the two standard local sections of the Hopf fibration $\tilde{\Sigma} \to S$. The transition function relating these sections takes values in $G(\mathbf{J}_0) \subset G$. The local sections are also local sections for $\tilde{S}^5 \to S$ and as such have the same transition function. Restricted to the equator the transition function represents the nontrivial generator of the fundamental group of $G = SO(3)$ and hence $\tilde{S}^5 \to S$ is the nontrivial bundle.

To prove the facts regarding Z observe that $\tilde{S}^5/G(\mathbf{J}_0)$ is isomorphic to the associated bundle $\tilde{S}^5 \times_G (G/G(\mathbf{J}_0))$.

Proofs of theorems 2 and 3. Any triangle can be made to lie in the xy plane by a rotation so that Q_{planar} is a slice for the G action on Q. An oriented triangle can be made planar in a unique way, up to rotation. An unoriented triangle can be made planar in two rotationally inequivalent ways, the two ways being related by reflection. In other words: $Q/G = Q_{planar}/O(2)$, whereas $\tilde{Q}/G = Q_{planar}/SO(2)$. This accounts for the difference in the two quotients. These two identifications are isometries, since Q_{planar} is totally geodesic. The last space is C^2/S^1 which is isometric to the cone over the sphere $S^2(\frac{1}{2})$. The quotient group $O(2)/SO(2)$ is the two-element group and accounts for the branched cover $S \to S_+$. The derivation of the formula for the normalized height z_1 can be found in Hsiang [3] and in our appendix.

The action by homotheties commutes with rotations so it descends to the quotient where it remains a dilation: $d(\lambda a, \lambda b) = \lambda d(a, b)$. (Here a, b represent similarity classes and d is the distance function.) Since I is homogeneous of degree 2, and since S^5 is defined by $I = 1$, the dilation parameter λ equals $\sqrt{I}$.

Proof of theorem 4. It follows from the discussion of the previous section, the above proofs and the fact that any curve of planar triangles has angular momentum in the $\hat{e}_3$ direction, that the restriction of $\mathbf{A}$ to $\Sigma \subset Q$ is $\Gamma\hat{e}_3$ where Γ is the canonical connection for the Hopf fibration. One can choose a local section s: $U \subset S \to \Sigma$ for the Hopf fibration such that $s^*\Gamma = -\frac{1}{2}z_1 \mathrm{d}\theta_1$. (The domain of this section is the sphere minus a 'branch cut' – a geodesic arc connecting the north and south pole.) It follows that

$$s^*\mathbf{A} = -(\tfrac{1}{2}z_1 \mathrm{d}\theta_1)\hat{e}_3. \tag{8}$$

(An alternative proof is given in appendix B.)

Remark. The principal axis frame discussed above provides the correct choice of local section. This is shown in Appendix B. Set $\psi_2 \cong 0 \cong \theta_2$ in the calculations there to obtain the above form of $s^*\mathbf{A}$.

Now $\mathbf{A}$ is 'basic' with respect to the action of the group $\mathbb{R}^+$ of dilations. This means that

$$\sigma_\lambda^* \mathbf{A} = \mathbf{A} \tag{9}$$

$$\frac{\partial}{\partial \lambda} \rfloor \mathbf{A} = 0 \tag{10}$$

where $\sigma_\lambda : \tilde{Q} \to \tilde{Q}$ is homothety by $\lambda \in \mathbb{R}^+$ and $\frac{\partial}{\partial\lambda}\rfloor$ denotes inner product with the infinitesimal generator $\frac{\partial}{\partial\lambda}$ of homotheties. (To see (9) observe that $\mathbb{I}(\lambda q) = \lambda^2 \mathbb{I}(q)$ and $\sigma_\lambda^* J = \Sigma m_a \lambda \mathbf{q}_a \times d(\lambda \mathbf{q}_1) = \lambda^2 \mathbf{J}$, and use the definition $\mathbf{A} = \mathbb{I}^{-1}\mathbf{J}$. To see (10) observe that the angular momentum of a pure dilational motion is 0 which means that $\frac{\partial}{\partial\lambda}\rfloor \mathbf{A} = 0$.) Extend the section s by making it constant under homothety. Then, by the homothety invariances of $\mathbf{A}$, formula (8) still holds for this extended section. Let $\mathbf{U}$ denote the local frame induced by s. It is a map to $G = SO(3)$ defined by writing $\tilde{q} = \mathbf{U}(\tilde{q}) s(\pi(\tilde{q}))$. The induced local trivialization of our principal bundle $\tilde{Q}^* \to \tilde{Q}^*/G$ is then $\tilde{q} \mapsto (\pi(\tilde{q}), \mathbf{U}(\tilde{q}))$. Note that $\mathbf{U}(\tilde{q})_3 = \mathbf{U}(\tilde{q})(e_3) = \mathbf{n}$, is the normal vector of the oriented triangle $\tilde{q}$. Using this fact, the transformation formula for connections, and the fact that under our identification of the Lie algebra of G with $\mathbb{R}^3$ the adjoint action of G becomes its usual action on $\mathbb{R}^3$, we see that with respect to our local trivialization we have:

$$\mathbf{A} = -(\tfrac{1}{2} z_1 \mathrm{d}\theta_1)\mathbf{n} + (\mathrm{d}U)U^{-1}$$

where $(\mathrm{d}U)U^{-1} = \Theta$ denotes the pull-back of the Maurer–Cartan form on G by the map U.

We now have

$$\alpha = -\frac{1}{2}(z_1 d\theta_1) z_2 + \mathbf{e}_3 \cdot \Theta,$$

since $\mathbf{e}_3 \cdot \mathbf{n} = z_2$. It is well-known (see [10]) that the two-form $\mathrm{d}(\mathbf{e}_3 \cdot \Theta)$ pushes down to $S^2 = G/G(\mathbf{J}_0)$ and that this push-down is the area form $-\mathrm{d}z_2 \wedge \mathrm{d}\theta_2$. Thus $\mathrm{d}\alpha = -\{\frac{1}{2}\mathrm{d}(z_1 z_2) \wedge \mathrm{d}\theta_1 + \mathrm{d}z_2 \wedge \mathrm{d}\theta_2\}$, which is the claimed formula for Ω.

Remark. We think of the local frame $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3)$ as a moving frame attached to our triangle, chosen so that $\mathbf{U}_3 = \mathbf{n}$ is the triangle's normal. As discussed in the introduction, and shown in Appendix B, the specific frame which leads to our choice of 'gauge', i.e. $s^*\mathbf{A}$, is the principal axis guage. See also Hsiang [3], or Reinsch and Littlejohn [13]and references therein. Note that the principal axis eigenframe is not well-defined at the north and south poles since these points correspond to weighted triangles for which $\mathbb{I}$ has double eigenvalues. (If the masses are all equal these are the equilateral triangles.) A branch cut from the north to south pole is also necessary, for if we traverse a small loop encircling the fibre over one of the poles then we wil find that the principal axis frame rotates by 180 degrees, not by 360 degrees.

4. Derivation of the reconstruction formula

In this section we derive our reconstruction formula, (2).

4.1. Closing the loop: a loop and a disc in Z

Let

$$s(t) = \pi(\tilde{q}(t)) \in S$$

be the curve of similarity classes represented by our three-body motion. It is a closed curve on the base two-sphere. Let

$$c_{\mathbf{J}_0}(t) = \pi_{\mathbf{J}_0}(\tilde{q}(t)) \in Z$$

denote the projection of $\tilde{q}(t)$ to Z. Although $s(t)$ is closed, the reduced curve $c_{\mathbf{J}_0}(t)$ need not be. There is a canonical way to close it. To see this, observe that $Z \to S$ has two canonical sections. One consists of equivalence classes of triangles whose normals are pointing along the $\mathbf{J}_0$ axis, and the other consists of those whose normals are antiparallel to the $\mathbf{J}_0$ axis. We will call these sections, or their values at a particular similarity class, the 'north' and 'south' poles. Since $s(t)$ is closed, both endpoints of the curve $c_{\mathbf{J}_0}(t)$ lie on the same spherical fibre over $s(0) = s(t_1)$. On this fixed fibre draw the two geodesic arcs from the north pole to $c_{\mathbf{J}_0}(0)$ and from $c_{\mathbf{J}_0}(t_1)$ back, then sandwich the reduced curve $c_{\mathbf{J}_0}(t)$ in between. The resulting closed curve will be denoted $\gamma_{\mathbf{J}_0}$. It is the projection of a closed curve $\gamma(t)$ in $\tilde{Q}$. See the figure.

Finally, Z is simply connected so that $\gamma_{\mathbf{J}_0}$ bounds some disc, $D \subset Z$.

4.2. The loop in Q

We now construct the loop γ in $\tilde{Q}$ over which we integrate. Its projection to Z is the loop just described above. The loop γ is obtained by concatenating the dynamic curve $\tilde{q}(t)$ with several group orbits, denoted $c_i(t)$ or $h(t)$. The act of concatenating two curves, one ending where the other begins, is defined in the obvious manner, and will be denoted by '$*$' below.

To construct the group curves we will use the following exponential notation for rotations. If $\mathbf{v} \in \mathbb{R}^3$ then $\exp(\mathbf{v})$, will mean the counter-clockwise rotation about the axis spanned by $\mathbf{v}$ by $\|\mathbf{v}\|$ radians. This is the standard Lie theoretic exponential map if we use the standard identification of $\mathbb{R}^3$ with the Lie algebra of the rotation group. If $\mathbf{v}$ is a unit vector, and $\theta \in \mathbb{R}$ then $\exp(\theta\mathbf{v})$ is a rotation by θ radians about the $\mathbf{v}$ axis. Let $\mathbf{n}_0$ and $\mathbf{n}_1$ be the initial and final normal vectors to our curve of oriented triangles, as in assumption (2), above. Form unit vectors

$$\xi_0 = \frac{1}{\|\mathbf{J}_0 \times \mathbf{n}_0\|}\mathbf{n}_0 \times \mathbf{J}_0$$

and

$$\xi_1 = \frac{1}{\|\mathbf{n}_1 \times \mathbf{J}_0\|}\mathbf{J}_0 \times \mathbf{n}_1$$

and corresponding one-parameter subgroups $\exp(s\xi_0)$, $\exp(s\xi_1)$ of rotations. Then:

$$R_0 = \exp(\phi_2(0)\xi_0)$$

and

$$R_1 = \exp(\phi_2(t_1)\xi_1)$$

where R_0, R_1 are the rotation matrices of our reconstruction formula (1), ϕ_2 is the angle in our parameterization of Z, and $\phi_2(t)$ is its value along the reduced curve $\gamma_{\mathbf{J}_0}(t)$: $J\cos(\phi_2(t)) = \mathbf{J}_0 \cdot \mathbf{n}(t)$.

Let $\tilde{q}(t)$ be the oriented three-body motion. Consider the concatenation

$$\gamma = c_0 * \tilde{q} * c_1 * c_{\mathbf{J}_0} * h$$

of the following curves:

$$\begin{aligned} c_0(t) &= \exp(-t\xi_0)R_0\tilde{q}(0) \qquad 0 \leqslant t \leqslant \phi_2(0) \qquad \tilde{q}(t)0 \leqslant t \leqslant t_1 \\ c_1(t) &= \exp(t\xi_1)\tilde{q}(t_1) \qquad 0 \leqslant t \leqslant \phi_2(t_1) \\ c_{\mathbf{J}_0}(t) &= \exp(-t\frac{1}{J_0}\mathbf{J}_0)R_1\tilde{q}(t_1) \qquad 0 \leqslant t \leqslant \Delta\theta \end{aligned}$$

and

$$h(t) = e^{at}c_J(\Delta\theta).$$

The interval of definitions of the curves are chosen so that the endpoint of one curve is the initial point of the next and so the concatenations are well-defined. The constant a and the time of stopping for the final purely dilational curve $h(t)$ are chosen so that its endpoint is the beginnning point, $R_0\tilde{q}(0)$, for $\gamma(t)$.

4.3. Line integrals

We have:

$$\int_\gamma \alpha = \int_{c_1} \alpha + \int_{\tilde{q}} \alpha + \int_{c_2} \alpha + \int_{c_J} \alpha + \int_h \alpha$$

CLAIM:

$$\begin{aligned} \int_{c_0} \alpha &= 0 \\ \int_{c_1} \alpha &= 0 \\ \int_{\tilde{q}} \alpha &= \int_0^{t_1} \omega(t)\mathrm{d}t \\ \int_{c_{\mathbf{J}_0}} \alpha &= -\Delta\theta \\ \int_h \alpha &= 0. \end{aligned}$$

The first two integrals vanish because

$$c^*\alpha = \mathbf{J}_0 \cdot c^*\mathbf{A} = \mathbf{J}_0 \cdot \omega \mathrm{d}s$$

whenever $c(s) = \exp(s\omega)c(0)$ is the orbit generated by a one-parameter subgroup of G. (This follows immediately from one of the defining properties of connections.) Now use the fact that the infinitesimal generators $\omega = \xi_0, \xi_1$ for the curves c_0, c_1 are perpendicular to $\mathbf{J}_0$. To evaluate the third integral, observe that $c_{\mathbf{J}_0}$ is also the orbit of a one-parameter subgroup, but its generator is the unit vector along $\mathbf{J}_0$. The vanishing of the integral over the homothety path h follows immediately from the homothety invariance of the connection already discussed. The integrand for the dynamic path $\tilde{q}(t)$ was already discussed. See remark 1 near the end of subsection 2.6. where we noted that $q^*\alpha = \omega \mathrm{d}t$.

An application of Stokes' theorem and the formulae relating $\mathrm{d}\alpha$ to Ω now prove our reconstruction formula, (2).

Appendix A. Identification of the shape sphere

Following the discusson of subsection 3.1 it suffices to understand the geometry of the space of similarity classes of weighted triangles for the planar three-body problem. We identify the plane in which the bodies move with the complex plane. Then we replace the spatial configuration space Q above by the planar configuration space Q_{planar} of triples $q = (q_1, q_2, q_3)$ of complex numbers, subject to the constraint $\Sigma_a m_a q_a = 0$. The space of similarity classes of triangles in the plane forms a two-sphere, and a three-body motion describes a curve $w(t)$ on this sphere.

Let us describe the sphere S of similarity classes explicitly. First, we diagonalize the mass matrix (kinetic energy) by introducting Jacobi coordinates

$$\xi_1 = q_1 - q_3$$

and

$$\xi_2 = \frac{-m_1}{m_1 + m_3} q_1 + q_2 + \frac{-m_3}{m_1 + m_3} q_3,$$

and normalized Jacobi coordinates

$$\zeta_1 = \sqrt{\mu_1}\xi_1,$$

and

$$\zeta_1 = \sqrt{\mu_2}\xi_2,$$

where the reduced masses μ_i are defined by $\frac{1}{\mu_1} = \frac{1}{m_1} + \frac{1}{m_3}$ and $\frac{1}{\mu_2} = \frac{1}{m_1+m_3} + \frac{1}{m_2}$. Then

$$\Sigma m_a \|\dot{q}_a\|^2 = \|\dot{\zeta}_1\|^2 + \|\dot{\zeta}_2\|^2.$$

Rotations by an angle θ induce the transformation $(\zeta_1, \zeta_2) \mapsto (\mathrm{e}^{\mathrm{i}\theta}\zeta_1, \mathrm{e}^{\mathrm{i}\theta}\zeta_1)$ It follows that the vector

$$\mathbf{w} = (w_1, w_2, w_3)$$

defined by

$$w_1 = \frac{1}{2}(\|\zeta_1\|^2 - \|\zeta_2\|^2),$$

and

$$w_2 + \mathrm{i}w_3 = \zeta_1\bar{\zeta}_2$$

is invariant under rotations. We calculate that the height coordinate w_3 is

$$w_3 = 2\sqrt{\frac{m_1 m_2 m_3}{m_1 + m_2 + m_3}} A$$

where A is the (oriented) area of the triangle q. In other words,the height on S represents the triangle's area, A.

The sphere of radius $\frac{1}{2}$

$$S = \{\mathbf{w} : w_1^2 + w_2^2 + w_3^2 = \tfrac{1}{4}\} = \{\mathbf{w} : I = 1\}$$

and is naturally identified with the space of similarity classes of planar triangles. Functions on S can be thought of as functions in the w_i which are homogeneous of degree 0 with respect to dilations. As such the height w_3 on S is the coordinate $\frac{2}{I}\sqrt{\frac{m_1 m_2 m_3}{m_1+m_2+m_3}} A$. The normalized height used in the body of the text is related to this coordinate by $w_3 = \frac{1}{2} z_1$, so that $z_1 = \frac{4}{I}\sqrt{\frac{m_1 m_2 m_3}{m_1+m_2+m_3}}$.

The w_i can also be easily written for oriented triangles in space. Using the same formulae for the Jacobi vectors ζ_1, ζ_1. Then w_1 is given by the same formula as before.

$$w_2 = \zeta_1 \cdot \zeta_2$$

and

$$w_3 = \mathbf{n} \cdot (\zeta_1 \times \zeta_2).$$

Appendix B. Principal axis calculations

Here we directly calculate the connection form $\mathbf{A}$ and its third component α, using the eigenframes for the inertia tensor. This gives an alternative proof of the formula (4) form Ω.

Let $\mathbf{u}, \mathbf{v}$ be normalized Jacobi vectors, so that $I = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$. The angular momentum is given by

$$\mathbf{J} = \mathbf{u} \times \mathrm{d}\mathbf{u} + \mathbf{v} \times \mathrm{d}\mathbf{v}.$$

Let $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ be a right-handed basis of eigenvectors for the inertia tensor $\mathbb{I}$, with $\mathbf{f}_3$ chosen to be the normal vector to the plane defined by the three bodies. Then

$$\mathbb{I}\mathbf{f}_1 = I(\sin(\psi/2))^2 f_1$$

$$\mathbb{I} f_2 = I(\cos(\psi/2))^2 f_2$$

$$\mathbb{I} f_3 = I f_3.$$

for some angle ψ. For clarity, we will write these first two eigenvalues also as λ_1, λ_2; thus

$$\lambda_1 = I(\sin(\psi/2))^2 \ , \ \lambda_2 = I(\cos(\psi/2))^2.$$

Since f_3 is the normal vector, we have

$$\mathbf{u} = A\mathbf{f}_1 + B\mathbf{f}_2$$

$$\mathbf{v} = C\mathbf{f}_1 + D\mathbf{f}_2.$$

Now $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = I = A^2 + B^2 + C^2 + D^2$. Setting

$$A = R\cos(\theta/2)\cos(\psi/2) \ , \ B = -R\sin(\theta/2)\sin(\psi/2)$$

$$C = R\sin(\theta/2)\cos(\psi/2) \ , \ D = R\cos(\theta/2)\sin(\psi/2)$$

where $R^2 = I$ guarantees that $\mathbf{f}_1, \mathbf{f}_2$ are indeed eigenvectors for the moment of inertia tensor $\mathbb{I} = \mathbb{I}(\mathbf{u}, \mathbf{v})$. Also one checks that the eigenvalues λ_1, λ_2 of $\mathbb{I}$ are indeed as given. A direct calculation using the Hopf coordinates (see e.g. appendix A) shows that θ, ψ are spherical coordinates on the shape sphere, with $z_1 = \sin(\psi)$. When the frame $\mathbf{f}_i$ is taken to be the coordinate basis, the configuration $\mathbf{u}, \mathbf{v}$ is planar, and the choices made define a local section of $\tilde{Q} \to \tilde{Q}/G$, the 'principal axis gauge'. This is the choice of gauge (local section) used for the formulae in the text.

We now expand the eigenframe in terms of an inertia-fixed frame $\mathbf{e} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. It will suffice to assume that the normal vector $\mathbf{n} = \mathbf{f}_3$ lies in the 1–3 plane. Thus:

$$\mathbf{f}_1 = \cos(\theta_2)\sin(\psi_2)\mathbf{e}_1 - \sin(\theta_2)\mathbf{e}_2 + \cos(\theta_2)\cos(\psi_2)\mathbf{e}_3$$
$$\mathbf{f}_2 = \sin(\theta_2)\sin(\psi_2)\mathbf{e}_1 + \cos(\theta_2)\mathbf{e}_2 + \sin(\theta_2)\cos(\psi_2)\mathbf{e}_3$$
$$\mathbf{f}_3 = -\cos(\psi_2)\mathbf{e}_1 + \sin(\psi_2)\mathbf{e}_3$$

and we express

$$\mathbf{e}_3 = \alpha\mathbf{f}_1 + \beta\mathbf{f}_2 + \gamma\mathbf{f}_3.$$

with

$$\alpha = \cos(\theta_2)\cos(\psi_2) \qquad \beta = \sin(\theta_2)\cos(\psi_2) \qquad \gamma = \sin(\psi_2).$$

Warning: the angle ψ_2 is not the angle ϕ_2 of the text. It is chosen so that $z_2 = \sin(\psi_2)$.

Differentiating the eigenframe we find:

$$\begin{pmatrix} \mathrm{d}f_1 \\ \mathrm{d}f_1 \\ \mathrm{d}f_3 \end{pmatrix} = \Omega \begin{pmatrix} f_1 \\ f_1 \\ f_3 \end{pmatrix}$$

where Ω is the skew-symmetric matrix of one-forms:

$$\Omega = \begin{pmatrix} 0 & \Omega_3 & -\Omega_2 \\ -\Omega_3 & 0 & \Omega_1 \\ \Omega_2 & -\Omega_1 & 0 \end{pmatrix}$$

with entries

$$\Omega_1 = -\sin(\theta_2)\mathrm{d}\psi_2 \qquad \Omega_2 = +\cos(\theta_2)\mathrm{d}\psi_2 \qquad \Omega_3 = -\mathrm{d}\theta_2.$$

For the rest of the calculation we will suppose that $R = I = 1$. Homothety-invariance will yield the general formula for our forms from the formulas with this restriction.

With $R = 1$ we have

$$\mathrm{d}\mathbf{u} = \mathrm{d}A\mathbf{f}_1 + \mathrm{d}B\mathbf{f}_2 + A(\Omega_3\mathbf{f}_2 - \Omega_2\mathbf{f}_3) + B(-\Omega_3\mathbf{f}_1 - \Omega_1\mathbf{f}_3).$$

Then:

$$\mathbf{u} \times \mathrm{d}\mathbf{u} = (-A\Omega_2 + B\Omega_1)(B\mathbf{f}_1 - A\mathbf{f}_2) + \{A(\mathrm{d}B + A\Omega_3) - B(\mathrm{d}A - B\Omega_3)\}\mathbf{f}_3.$$

Similarly:

$$\mathbf{v} \times \mathrm{d}\mathbf{v} = (-C\Omega_2 + D\Omega_1)(D\mathbf{f}_1 - C\mathbf{f}_2) + \{C(\mathrm{d}D + C\Omega_3) - D(\mathrm{d}C - D\Omega_3)\}\mathbf{f}_3.$$

We can now directly calculate the connection-form $\mathbf{A} = \mathbb{I}^{-1}\mathbf{J}$.

$$\begin{aligned} \mathbb{I}^{-1}\mathbf{J} = {} & \frac{1}{\lambda_1}\{(-A\Omega_2 + B\Omega_1)B + (-C\Omega_2 + D\Omega_1)D\}\mathbf{f}_1 \\ & + \frac{1}{\lambda_2}\{(-A\Omega_2 + B\Omega_1)(-A) + (-C\Omega_2 + D\Omega_1)(-C)\}\mathbf{f}_2 \\ & + \frac{1}{1}\{A(\mathrm{d}B + A\Omega_3) - B(\mathrm{d}A - B\Omega_3) + C(\mathrm{d}D + C\Omega_3) - D(\mathrm{d}C - D\Omega_3)\}\mathbf{f}_3. \end{aligned}$$

After some algebra we find:

$$\mathbf{A} = -(\sin(\theta_2)\mathrm{d}\psi_2)\mathbf{f}_1 + (\cos(\theta_2)\mathrm{d}\psi_2)\mathbf{f}_2 - (\tfrac{1}{2}z_1\mathrm{d}\theta_1 + \mathrm{d}\theta_2)\mathbf{f}_3.$$

Then $\alpha = -\gamma(\frac{1}{2}z_1\mathrm{d}\theta_1 + \mathrm{d}\theta_2) = -(\frac{1}{2}z_2z_1\mathrm{d}\theta_1 + z_2\mathrm{d}\theta_2)$, and as desired $\Omega = \mathrm{d}\alpha$ is given by the claimed formula (4) of the text.

Acknowledgments

I would like to thank Wu-Yi Hsiang for suggesting this problem, and for helpful conversations. I would also like to thank Mathias Reinsch (Physics, UC Berkeley) for several suggestions, and for a careful reading of this paper in which he corrected numerous errors.

References

[1] Guichardet A 1984 *Ann. Inst. H. Poincare, Phys. Theor.* **40** 329

[2] Guillemin V and Sternberg S 1984 *Symplectic Techniques in Physics* (Cambridge: Cambridge University Press)

[3] Hsiang W-Y 1994 Geometric study of the three-body problem, I *UC Berkeley preprint PAM-620*

[4] Iwai T 1987 *Ann. Inst. Henri Poincare* **47** 199

[5] Levi M 1993 Geometric Phases in the Motion of Rigid Bodies *Archive for Rational Mechanics and Analysis* **122** 213

[6] Marsden J, Montgomery R and Ratiu T 1990 Reduction, Symmetry, and Berry's Phase in Mechanics *Memoirs of the AMS* **88**

[7] Marsden J and Weinstein A 1974 *Rep. Math. Phys.* **5** 121

[8] Moeckel R 1988 *Proc. Hamiltonian Dynamical Systems, AMS Contemporary Math* ed K Meyer and C Saari **81** 1 (Providence, RI: American Mathematical Society)

[9] Montgomery R 1991 Appendix to: Optimal Control of Deformable Bodies and its Relation to Gauge Theory *Proc. Geometry of Hamiltonian Systems Workshop (1988)* ed T Ratiu (Berlin: Springer)

[10] Montgomery R 1991 *Am. J. Phys.* **59** 394

[11] Montgomery R 1984 *Lett. Math. Phys.* **8** 59

[12] Montgomery R 1986 Bundle Theory in Mechanics *PhD thesis*

[13] Reinsch M and Littlejohn R 1996 Gauge fields in the separation of rotations and internal motions in the n-body problem *Rev. Mod. Phys* to appear

[14] Saari D 1984 *Celestial Mechanics* **33** 299

[15] Saari D 1988 *Proc. Hamiltonian Dynamical Systems AMS Contemporary Math* ed K Meyer and C Saari **81** 23 (Providence, RI: American Mathematical Society)

[16] 1989 Shapere A and Wilczek F *Geometric Phases in Physics* (Singapore: World Scientific)

[17] Smale S 1970 *Inventiones* **10** 305; **11** 45

[18] Weinstein A 1978 *Lett. Math. Phys.* **2** 417