

# MINIMAL ENERGY CONFIGURATIONS OF GRAVITATIONALLY INTERACTING RIGID BODIES

RICHARD MOECKEL

ABSTRACT. Consider a collection of  $n$  rigid, massive bodies interacting according to their mutual gravitational attraction. A *relative equilibrium* motion is one where the entire configuration rotates rigidly and uniformly about a fixed axis in  $\mathbb{R}^3$ . Such a motion is possible only for special positions and orientations of the bodies. A *minimal energy* motion is one which has the minimum possible energy in its fixed angular momentum level. While every minimal energy motion is a relative equilibrium motion, the main result here is that a relative equilibrium motion of  $n \geq 3$  disjoint rigid bodies is never an energy minimizer. This generalizes a known result about point masses to the case of rigid bodies.

## 1. INTRODUCTION

The full  $n$ -body problem studies the motion of  $n$  rigid, massive bodies in  $\mathbb{R}^3$  moving under the influence of their mutual gravitational attraction. The usual  $n$ -body problem deals with point masses and provides a good model for celestial mechanics when the masses are far away from one another or are spherically symmetric. The full problem is especially important when asymmetrical masses interact at comparatively close range. In that case, tidal forces and other dissipative effects can lead to changes in the orbits and the rotational motions. Dissipative forces due to tidal interactions among the bodies lead to a decrease in the total energy of the system, but leave the total angular momentum unchanged. From this point of view it is interesting to ask for the minimal energy states for a given level of angular momentum.

Fixing the angular momentum and center of mass gives a submanifold of the phase space. For the point mass  $n$ -body problem, it has long been known that the critical points of the energy on such a momentum level are the relative equilibrium states [8, 9]. The same holds true for the full  $n$ -body problem. This means that the entire configuration rotates uniformly around some axis in  $\mathbb{R}^3$ . The centers of mass move on circles around the axis and the rigid bodies rotate simultaneously to maintain phase locking. Pluto and its moon Charon provide a rough example for  $n = 2$ .

---

*Date:* August 18, 2016, preliminary version.

*2000 Mathematics Subject Classification.* 70F10, 70F15, 37N05, 70G40, 70G60.

*Key words and phrases.* Celestial mechanics, rigid bodies.

Research supported by NSF grant DMS-1208908 .

If such a motion is to arise due to energy dissipation, it should be a local minimum of the energy and not just a critical point. While such energy minimizing motions are possible for  $n = 1, 2$ , it will be shown below that they are impossible for  $n \geq 3$ . The implication for celestial mechanics is that starting with  $n \geq 3$  bodies, one expects that dissipative effects will lead to the collisions of some of the masses so that in the end they form one or two amalgamated bodies or else will result in some of the bodies moving off to infinity.

The fact that relative equilibria cannot be energy minimizers was known for the point mass case [5]. It was conjectured for the full  $n$ -body problem by Scheeres [6] and this paper was written specifically to settle this conjecture. In light of this result, it is interesting to look for energy minimizers among motions where the bodies are in contact and Scheeres has done this in the case of a few spherical bodies.

In addition to proving the main result in Theorem 4, we also provide elementary proofs of some known facts about relative equilibria. Namely, relative equilibria in phase space coincide with critical points of the energy on manifolds of fixed angular momentum (Theorem 1) and relative equilibrium configurations can be viewed as critical points of an *amended potential* function on configuration space with corresponding local minima (Theorem 2). These are special cases of general facts about relative equilibria for mechanical systems with symmetry as described, for example, in [8, 1, 4, 7] but a more elementary approach might be of some value. In addition to the amended potential, we also work with another, simpler function used extensively by Scheeres. This function has the same critical points as the amended potential (Theorem 3) and, at least under certain conditions, they also have the same local minima. This is used in the proof of Theorem 4.

## 2. EQUATIONS OF MOTION

Consider a collection of  $n$  rigid, massive bodies in  $\mathbb{R}^3$ . Each body can be described in its own body coordinate system by a compact subset  $\mathcal{B}_i \subset \mathbb{R}^3$  together with a mass measure  $dm_i$  on  $\mathcal{B}_i$ ,  $i = 1, \dots, n$ . This might take the form  $dm_i = \nu_i(Q_i) dQ_i$  where  $\nu_i \geq 0$  is a continuous mass density function but other measures are also allowed provided all of the integrations which occur below are valid. Denote the  $i$ -th body coordinate system by  $Q_i \in \mathbb{R}^3$ . Then the total mass of the  $i$ -th body is given by the triple integral

$$m_i = \int_{\mathcal{B}_i} dm_i$$

and we assume  $m_i > 0$ . It is convenient to assume that its center of mass is at the origin in body coordinates, i.e.,

$$\int_{\mathcal{B}_i} Q_i dm_i = 0.$$

We will need the symmetric  $3 \times 3$  *inertia matrix* of  $\mathcal{B}_i$

$$(1) \quad I_i = \int_{\mathcal{B}_i} (|Q_i|^2 \mathbb{I} - Q_i Q_i^T) dm_i$$

where  $\mathbb{I}$  is the  $3 \times 3$  identity matrix. To avoid degenerate situations we will assume that the mass distributions are such that the matrices  $I_i$  are all invertible. This excludes point masses and one-dimensional mass distributions.

The position and orientation of the body with respect to the inertial coordinates,  $x \in \mathbb{R}^3$  is given by a time-dependent Euclidean transformation  $E_i(t)$  where

$$(2) \quad x(t, Q_i) = E_i(t)(Q_i) = A_i(t)Q_i + q_i(t) \quad Q_i \in \mathcal{B}_i.$$

The rotation matrix  $A_i(t) \in \mathbf{SO}(3)$  describes the orientation of the body and  $q_i(t) \in \mathbb{R}^3$  is the center of mass in the inertial system.

The positions and orientations of all of the bodies is given by  $Z = (q_1, \dots, q_n, A_1, \dots, A_n) \in \mathbb{R}^{3n} \times \mathbf{SO}(3)^n$ . The *configuration space* will be the open subset of  $\mathbb{R}^{3n} \times \mathbf{SO}(3)^n$  where the bodies are disjoint

$$\tilde{\mathcal{U}} = \{Z : E_i(\mathcal{B}_i) \cap E_j(\mathcal{B}_j) = \emptyset, i \neq j\}.$$

The gravitational interaction is governed by the *Newtonian potential* function. For each pair of indices  $(i, j)$ ,  $i \neq j$ , there is a mutual potential

$$U_{ij}(q_i, q_j, A_i, A_j) = \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{dm_i dm_j}{|q_i - q_j + A_i Q_i - A_j Q_j|}$$

which involves integrals over each body. The Newtonian potential is given by

$$U(Z) = \sum_{i < j} U_{ij}.$$

This is a well-defined, smooth, positive function  $U : \tilde{\mathcal{U}} \rightarrow \mathbb{R}$ . Although we are calling  $U(Z)$  the Newtonian potential, the potential energy of the system is  $-U(Z)$ .

The velocity of the point (2) is

$$\dot{x}(t, Q_i) = \dot{q}_i(t) + \dot{A}_i(t)Q_i = v_i(t) + A_i(t)\hat{\Omega}_i(t)Q_i$$

where  $v_i$  denotes the velocity of the center of mass and

$$\hat{\Omega}_i(t) = A_i^{-1}(t)\dot{A}_i(t)$$

is the antisymmetric angular velocity matrix with respect to body coordinates. We will also make use of the corresponding *angular velocity vector*  $\Omega_i \in \mathbb{R}^3$  such that  $\hat{\Omega}_i u = \Omega_i \times u$  for all vectors  $u \in \mathbb{R}^3$ . If  $\Omega_i(t)$  is known then the rotation matrix  $A_i(t)$  can be reconstructed from the differential equation

$$\dot{A}_i(t) = A_i(t)\hat{\Omega}_i(t).$$

In addition to the configuration variables  $q_i, A_i$  we will use  $v_i, \Omega_i$  as velocity variables on the phase space  $T\tilde{\mathcal{U}}$ .

To find the equations of motion, we will consider the translational and rotational motions of  $\mathcal{B}_i$  separately. The motion of the centers of mass  $q_i$  are governed by

$$m_i \ddot{q}_i(t) = m_i \dot{v}_i(t) = f_i(Z)$$

where  $f_i(Z)$  is the total force on  $\mathcal{B}_i$  due to the other bodies. The force vector acting at the point (2) due to the other bodies is given by

$$(3) \quad g_i(Z, Q_i) = - \sum_{j \neq i} \int_{\mathcal{B}_j} \frac{(q_i - q_j + A_i Q_i - A_j Q_j) dm_j}{|q_i - q_j + A_i Q_i - A_j Q_j|^3}.$$

Integrating this over  $\mathcal{B}_i$  gives

$$f_i(Z) = - \sum_{j \neq i} \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{(q_i - q_j + A_i Q_i - A_j Q_j) dm_i dm_j}{|q_i - q_j + A_i Q_i - A_j Q_j|^3} = U_{q_i}(Z).$$

Here  $U_{q_i}$  denotes the partial gradient vector with respect to  $q_i$ .

Later we will also need the Hessian quadratic form of matrix of  $U$  with respect to the  $q_i$  variables, which we call  $D_q^2 U$ . If  $w \in \mathbb{R}^{3n}$  is the vector such that  $w_i \in \mathbb{R}^3$  represents the displacement of  $q_i$  then

$$(4) \quad D_q^2 U(w, w) = \sum_{i < j} \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{dm_i dm_j}{r_{ij}^3} (-|w_{ij}|^2 + 3(u_{ij} \cdot w_{ij})^2)$$

where  $w_{ij} = w_i - w_j \in \mathbb{R}^3$  and where

$$r_{ij} = |q_i - q_j + A_i Q_i - A_j Q_j| \quad u_{ij} = \frac{q_i - q_j + A_i Q_i - A_j Q_j}{|q_i - q_j + A_i Q_i - A_j Q_j|}.$$

The rotational equations of motion are best described in terms of angular momenta and the inertia matrices. In the inertial frame, the angular momentum vector of the  $i$ -th body with respect to the origin is

$$\begin{aligned} \lambda_i &= \int_{\mathcal{B}_i} x(t, Q_i) \times \dot{x}(t, Q_i) dm_i = m_i q_i \times v_i + \int_{\mathcal{B}_i} A_i (Q_i \times (\Omega_i \times Q_i)) dm_i \\ &= m_i q_i \times v_i + A_i I_i \Omega_i. \end{aligned}$$

where  $I_i$  is the inertia matrix (1). Since we already have equations for the motion of the center of mass, we can focus on the angular momentum with respect to the center of mass

$$\mu_i(t) = A_i(t) I_i \Omega_i(t).$$

This satisfies the differential equation

$$\dot{\mu}_i(t) = \tau_i(Z)$$

where  $\tau_i(Z)$  is the total torque with respect to the center of mass on  $\mathcal{B}_i$  due to the other bodies. The torque vector acting at the point (2) due to the other bodies is given by the cross product

$$A_i(t) Q_i \times g_i(Z, Q_i)$$

where  $g_i(Z)$  is given by (3). Integrating this gives

$$\tau_i(Z) = - \sum_{j \neq i} \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{A_i Q_i \times (q_i - q_j + A_i Q_i - A_j Q_j) dm_i dm_j}{|q_i - q_j + A_i Q_i - A_j Q_j|^3}.$$

Pulling back to the body frame of  $\mathcal{B}_i$  using  $A_i^{-1} = A_i^T$  we get the body angular momentum vector

$$M_i(t) = I_i \Omega_i(t)$$

which satisfies the differential equation

$$\dot{M}_i = M_i \times \Omega_i + T_i$$

where

$$T_i(Z) = A_i^T \tau_i = - \sum_{j \neq i} \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{Q_i \times (A_i^T (q_i - q_j + A_i Q_i - A_j Q_j)) dm_i dm_j}{|q_i - q_j + A_i Q_i - A_j Q_j|^3}.$$

It is possible to interpret the torque vectors as derivatives of the Newtonian potential with respect to the rotation matrices  $A_i$ . To see this, note that a tangent vector to  $\mathbf{SO}(3)$  at the matrix  $A_i$  is represented by a curve of rotation matrices  $A_i R(t)$  where  $R(t) \in \mathbf{SO}(3)$  and  $R(0) = \mathbb{I}$ . The matrix  $\hat{\rho} = \dot{R}(0) \in \mathfrak{so}(3)$  can be identified with a vector  $\rho \in \mathbb{R}^3$  in the usual way via the cross-product. Let  $Z_i(t)$  be the curve in configuration space where  $A_i$  is replaced by  $A_i R(t)$  and all other variables are unchanged. Then after some computation we find

$$\frac{d}{dt} U(Z(t))|_{t=0} = T_i(Z) \cdot \rho.$$

Thus with these identifications, the torque vector  $T_i(Z)$  become a kind of partial gradient of  $U(Z)$  with respect to  $A_i$ . By abuse of notation we will write  $T_i(Z) = U_{A_i}(Z)$ . We will use this approach to handle differentiation with respect to the orthogonal matrices  $A_i$  throughout the paper.

Thus we have arrived at the equations of motion

$$(5) \quad \begin{aligned} m_i \ddot{q}_i &= f_i(Z) = U_{q_i}(Z) \\ \dot{M}_i &= M_i \times \Omega_i + T_i(Z) = M_i \times \Omega_i + U_{A_i}(Z) \\ \dot{A}_i &= A_i \hat{\Omega}_i \end{aligned}$$

with  $M_i = I_i \Omega_i$  and  $T_i$  the total torque on  $\mathcal{B}_i$  in the body frame. Since the inertia matrices  $I_i$  are invertible, these determine a system of first order differential equations on the phase space  $T\tilde{\mathcal{U}}$ .

These equations admit the usual symmetries and the corresponding constants of motion. First we have symmetry under translation of all of the bodies.  $q_i \mapsto q_i + c$ ,  $c \in \mathbb{R}^3$ , which leaves the potential  $U(Z)$  invariant. It follows that  $\sum_i U_{q_i}(Z) = 0$  and therefore the total momentum vector

$$p_{tot} = m_1 v_1 + \dots + m_n v_n$$

is constant. Without loss of generality we assume  $p_{tot} = 0$ . Then the center of mass is constant and may be taken as the origin of the inertial system. This amounts to restricting to a translation-reduced phase space  $T\mathcal{U}$  where

$$\mathcal{U} = \{Z \in \tilde{\mathcal{U}} : m_1 q_1 + \dots + m_n q_n = 0\}.$$

We have  $\dim \mathcal{U} = 6n - 3$  and  $\dim T\mathcal{U} = 12n - 6$ .

The problem is also symmetric under rotations. If  $R \in \mathbf{SO}(3)$  then the rotated configuration  $RZ$  has centers of mass  $Rq_i$  and orientation matrices  $RA_i$ ,  $i = 1, \dots, n$ . In other words  $\mathbf{SO}(3)$  acts on  $\mathbb{R}^{3n} \times \mathbf{SO}(3)^n$  diagonally from the left. The velocities of the centers of mass are also rotated to  $Rv_i$  but the body angular velocities  $\Omega_i$  are unchanged. As a result of the rotational symmetry, the total angular momentum vector in the inertial frame

$$(6) \quad \lambda = \sum_i m_i q_i \times v_i + \sum_i A_i I_i \Omega_i$$

is constant.

Finally the total energy

$$H(Z, \dot{Z}) = T(Z, \dot{Z}) - U(Z)$$

is constant, where  $T(Z, \dot{Z})$  is the kinetic energy

$$\begin{aligned} T(Z, \dot{Z}) &= \frac{1}{2} \sum_i m_i |v_i|^2 + \frac{1}{2} \sum_i \int_{\mathcal{B}_i} |\Omega_i \times Q_i|^2 dm_i \\ &= \frac{1}{2} \sum_i m_i |v_i|^2 + \frac{1}{2} \sum_i \Omega_i^T I_i \Omega_i. \end{aligned}$$

### 3. RELATIVE EQUILIBRIA

For a relative equilibrium motion, the configuration of  $n$  bodies rotates uniformly around a fixed axis through the origin in space. Let  $e \in \mathbb{R}^3$  be a unit vector specifying the direction of the rotation axis and let  $R(t) \in \mathbf{SO}(3)$  be the matrix with  $R(0) = \mathbb{I}$  representing rotation around the axis with constant angular speed  $\omega \neq 0$ . Suppose  $Z = (q_1, \dots, A_1, \dots) \in \mathcal{U}$  is the initial configuration of a relative equilibrium motion. Then

$$Z(t) = R(t)Z(0) = (q_1(t), \dots, A_1(t), \dots) = (R(t)q_1, \dots, R(t)A_1, \dots)$$

must be a solution of the equations of motion.

Since  $q_i(t) = R(t)q_i$  and since the angular velocity in the inertial frame is  $\omega e$ , we have

$$\begin{aligned} \dot{q}_i(t) &= \omega e \times q_i(t) = \omega R(t)(e \times q_i) \\ \ddot{q}_i(t) &= \omega^2 e \times (e \times q_i(t)) = \omega^2 R(t)(e \times (e \times q_i)). \end{aligned}$$

Rotation invariance of  $U(Z)$  implies that

$$U_{q_i}(Z(t)) = R(t)U_{q_i}(Z).$$

Substituting these formulas into the equations of motion shows that centers of mass of the relative equilibrium configuration  $Z$  must satisfy

$$(7) \quad U_{q_i}(Z) = \omega^2 m_i (e \times (e \times q_i)) = -\omega^2 m_i K_e q_i$$

where  $K_e$  is the projection onto the orthogonal plane  $e^\perp$ .

Similarly, from  $A_i(t) = R(t)A_i$  we find that the body angular velocity vector of  $\mathcal{B}_i$  is the constant vector

$$\Omega_i = \omega A_i^T e.$$

It follows that the body angular momentum vector  $M_i = I_i \Omega_i = \omega I_i A_i^T e$  must also be constant. On the other hand, the torque vector in body coordinates satisfies  $T_i(R(t)Z) = T_i(Z)$ , so the equations of motion give

$$0 = M_i \times \Omega_i + T_i(Z)$$

or

$$(8) \quad T_i(Z) + \omega^2 ((I_i A_i^T e) \times (A_i^T e)) = 0.$$

If  $Z \in \mathcal{U}$  satisfies (7) and (8), it will be called a *relative equilibrium configuration*. The point in the reduced phase space  $T\mathcal{U}$  with configuration variables  $Z$  and velocity variables

$$(9) \quad v_i = \omega e \times q_i \quad \Omega_i = \omega A_i^T e.$$

is the corresponding *relative equilibrium state*.

From equation (9) we find that the total angular momentum of a relative equilibrium state in the inertial frame is given by

$$\begin{aligned} \lambda_{re} &= \omega R(t) \left( \sum_i m_i q_i \times (e \times q_i) + \sum_i A_i I_i A_i^T e \right) \\ &= \omega R(t) I(Z) e \end{aligned}$$

where

$$(10) \quad I(Z) = \sum_i m_i (|q_i|^2 \mathbb{I} - q_i q_i^T) + \sum_i A_i I_i A_i^T$$

is the  $3 \times 3$  *total inertia matrix* of the whole configuration. Since  $\lambda$  is constant, the vector  $I(Z)e$  must be of the form  $ce$  for some constant  $c$ . In other words,  $e$  is an eigenvector of the total inertia tensor. Taking the inner product with  $e$  shows that the corresponding eigenvalue is  $c = G_e(Z)$ , where

$$(11) \quad G_e(Z) = e^T I(Z) e = \sum_i m_i q_i^T K_e q_i + \sum_i e^T A_i I_i A_i^T e$$

is the moment of inertia of the configuration  $Z$  with respect to the  $e$ -axis. So we have

$$(12) \quad \lambda_{re} = \omega I(Z) e = \omega G_e(Z) e.$$

Similarly, we find that the total energy of a relative equilibrium is

$$(13) \quad H_{re} = \frac{1}{2} G_e(Z) \omega^2 - U(Z).$$

In what follows we will be interested in relative equilibria with a given, nonzero angular momentum vector  $\lambda \in \mathbb{R}^3$ . Then the rotation axis and angular speed are uniquely determined by

$$(14) \quad e = \frac{\lambda}{|\lambda|} \quad \omega = \frac{|\lambda|}{G_e(Z)}.$$

A configuration  $Z \in \mathcal{U}$  admits a relative equilibrium motion with angular momentum  $\lambda$  if and only if it satisfies (7) and (8) with  $e, \omega$  given by (14), that is,

$$(15) \quad \begin{aligned} U_{q_i}(Z) + \frac{|\lambda|^2}{G_e(Z)^2} m_i K_e q_i &= 0 \\ T_i(Z) + \frac{|\lambda|^2}{G_e(Z)^2} ((I_i A_i^T e) \times (A_i^T e)) &= 0. \end{aligned}$$

In this case  $Z$  will be called a *relative equilibrium configuration for angular momentum*  $\lambda$ . The velocities are given by (9)

$$(16) \quad v_i = \frac{|\lambda|}{G_e(Z)} e \times q_i \quad \Omega_i = \frac{|\lambda|}{G_e(Z)} A_i^T e.$$

and the corresponding point in the phase space  $T\mathcal{U}$  will be called a *relative equilibrium state for angular momentum*  $\lambda$ . The energy of such a state is

$$(17) \quad H_\lambda = \frac{|\lambda|^2}{2G_e(Z)} - U(Z).$$

If  $Z$  is a relative equilibrium configuration for angular momentum  $\lambda$  and  $R \in \mathbf{SO}(3)$ , then  $RZ$  is a relative equilibrium configuration for angular momentum  $R\lambda$ . In particular, rotations which preserve  $\lambda$  also preserve the relative equilibria for  $\lambda$ . Thus every relative equilibrium is part of a circle of relative equilibria with the same angular momentum.

#### 4. MINIMAL ENERGY SOLUTIONS

Next we consider the problem of minimum energy states for a given value of the angular momentum vector. We will use the notation  $P = (Z, \dot{Z})$  to denote points of  $T\mathcal{U}$ . Fixing  $\lambda \neq 0$  determines an integral manifold  $\mathcal{M}_\lambda \subset T\mathcal{U}$ . We want to find states which locally or globally minimize the energy  $H = T(Z, \dot{Z}) - U(Z)$  on these manifolds.

**Lemma 1.** *For  $\lambda \neq 0$ ,  $\mathcal{M}_\lambda \subset T\mathcal{U}$  is a submanifold of codimension 3, that is,  $\dim \mathcal{M}_\lambda = 12n - 9$ .*

*Proof.* We need to show that the derivatives of the three components of  $\lambda$  together with the 6 linear equations defining  $T\mathcal{U}$  are linearly independent at every  $P \in \mathcal{M}_\lambda$ . Let  $\lambda_{q_i}, \lambda_{v_i}, \lambda_{\Omega_i}$  denote the  $3 \times 3$  matrices whose columns are the partial gradients of the three components  $\lambda$ . The analogous partial gradient matrices for the three components of  $m_1 q_1 + \dots + m_n q_n$ , and  $m_1 v_1 + \dots + m_n v_n$

$\dots + m_n v_n$ , are simply  $m_i \mathbb{I}$ . If the required linear independence did not hold, there would be vectors  $\alpha, \beta, \gamma \in \mathbb{R}^3$ , not all zero, such that

$$\lambda_{q_i} \alpha + m_i \beta = \lambda_{v_i} \alpha + m_i \gamma = \lambda_{\Omega_i} \alpha = 0$$

Furthermore, for every curve of matrices  $R(t) \in \mathbf{SO}(3)$  with  $R(0) = \mathbb{I}$ ,

$$\alpha \cdot \frac{d}{dt} \lambda(Z_i(t), \dot{Z})|_{t=0} \quad i = 1, \dots, n$$

where  $Z_i(t)$  is the curve of configurations where  $A_i$  is replaced by  $A_i(t) = A_i R(t)$  and all other variables are left constant. We will show that this can only happen when  $\lambda = 0$ .

The first two dependence conditions give

$$m_i v_i \times \alpha + m_i \beta = m_i \alpha \times q_i + m_i \gamma = 0.$$

We have  $(\sum_i m_i) \beta = -\sum_i m_i v_i \times \alpha = 0$  since the total momentum is zero. Thus  $\beta = 0$  and similarly  $\gamma = 0$ . Now the four dependence relations reduce to

$$\alpha \times q_i = v_i \times \alpha = I_i A_i^T \alpha = (I_i \Omega_i) \times (A_i^T \alpha) = 0.$$

This means that all of the vectors  $q_i, v_i, A_i I_i \Omega_i$  are scalar multiples of  $\alpha$  and, in addition, that  $I_i A_i^T \alpha = 0$ . It follows that

$$I_i A_i^T (A_i I_i \Omega_i) = I_i^2 \Omega_i = 0.$$

Since  $I_i$  is diagonalizable, it follows that  $I_i \Omega_i = 0$  too. All of this gives  $q_i \times v_i = A_i I_i \Omega_i = 0$  and so the angular momentum vector (6) is  $\lambda = 0$ .  $\square$

We are looking for local minima of the energy on  $\mathcal{M}_\lambda$  or, more generally, for critical points which are not necessarily local minima.

**Theorem 1.** *Let  $\lambda \in \mathbb{R}^3$  be any nonzero vector. A state  $P$  is a critical point of the restriction of the total energy function to  $\mathcal{M}_\lambda$  if and only if it is a relative equilibrium state.*

*Proof.* If  $P \in \mathcal{M}_\lambda$  has configuration  $Z$  and velocities  $v_i, \Omega_i$  and is a critical point of  $H$  restricted to  $\mathcal{M}_\lambda$  then there are vector Lagrange multipliers  $\alpha, \beta, \gamma \in \mathbb{R}^3$  such that

$$H_{q_i} = \lambda_{q_i} \alpha + m_i \beta \quad H_{v_i} = \lambda_{v_i} \alpha + m_i \gamma \quad H_{\Omega_i} = \lambda_{\Omega_i} \alpha$$

and such that for every curve of matrices  $R(t) \in \mathbf{SO}(3)$  with  $R(0) = \mathbb{I}$ ,

$$\alpha \cdot \frac{d}{dt} H(Z_i(t), \dot{Z})|_{t=0} = \alpha \cdot \frac{d}{dt} \lambda(Z_i(t), \dot{Z})|_{t=0} \quad i = 1, \dots, n$$

with  $Z_i(t)$  as above.

The first three conditions read

$$-U_{q_i} = m_i (v_i \times \alpha + \beta) \quad m_i v_i = m_i (\alpha \times q_i + \gamma) \quad I_i \Omega_i = I_i A_i^T \alpha$$

and the last one gives

$$-T_i(Z) = (I_i \Omega_i) \times (A_i^T \alpha).$$

As before we find that  $\beta = \gamma = 0$ . Then if we set  $\alpha = \omega e$  where  $e$  is a unit vector, the velocities are given by the relative equilibrium values (9) and the configuration variables satisfy (7) and (8). So we have a relative equilibrium state.

Conversely, if (7), (8) and (9) hold we get the critical point equations for  $H$  restricted to  $\mathcal{M}_\lambda$  with  $\alpha = \omega e$  and  $\beta = \gamma = 0$ .  $\square$

Theorem 1 characterizes the critical points  $P$  of the restriction of  $H(P)$  to  $\mathcal{M}_\lambda$  as relative equilibrium states. Next we will show that the configuration  $Z$  of such a critical point  $P$  must be a critical point of a function  $W_\lambda(Z)$ , the *amended potential*. Begin by fixing  $Z \in \mathcal{U}$  and  $\lambda \in \mathbb{R}^3$ . Then the angular momentum equation (6) defines an affine subspace of the velocity space  $T_Z\mathcal{U}$ :

$$\mathcal{S}_\lambda(Z) = \{v_i, \Omega_i : m_1 v_1 + \dots + m_n v_n = 0 \text{ and (6) holds}\}.$$

**Lemma 2.** *Fix  $Z \in \mathcal{U}$  and  $\lambda \neq 0$ . The equation  $\lambda = I(Z)\alpha$  has a unique solution  $\alpha(Z, \lambda) \in \mathbb{R}^3$  and then*

$$(18) \quad v_i = \alpha \times q_i \quad \Omega_i = A_i^T \alpha$$

are the velocities which minimize the energy over  $\mathcal{S}_\lambda(Z)$ . The minimum energy is given by the amended potential

$$(19) \quad W_\lambda(Z) = \frac{1}{2} \alpha(Z, \lambda)^T I(Z) \alpha(Z, \lambda) - U(Z) = \frac{1}{2} \lambda^T I(Z)^{-1} \lambda - U(Z).$$

*Proof.* The definition (10) shows that  $I(Z)$  is a sum of positive semi-definite  $3 \times 3$  matrices and all of the terms involving  $A_i$  are positive definite. It follows that  $I(Z)$  is positive definite and hence invertible. So  $\alpha(Z, \lambda) = I(Z)^{-1} \lambda$  is uniquely determined. Choosing the velocities as in (18) we find that the total momentum is zero and the angular momentum is  $I(Z)\alpha = \lambda$  so these velocities are in  $\mathcal{S}_\lambda(Z)$ .

To see that they give the minimum energy, note that the kinetic energy is a positive definite quadratic form in the velocities while the potential energy is constant on  $\mathcal{S}_\lambda(Z)$ . Viewing the kinetic energy as arising from an inner product on velocity space, it suffices to check that the vector (18) is orthogonal to the affine subspace  $\mathcal{S}_\lambda(Z)$ . The tangent space to  $\mathcal{S}_\lambda(Z)$  is the subspace consisting of velocities  $\tilde{v}_i, \tilde{\Omega}_i$  with  $m_1 \tilde{v}_1 + \dots + m_n \tilde{v}_n = 0$  and such that

$$\sum_i m_i q_i \times \tilde{v}_i + \sum_i A_i I_i \tilde{\Omega}_i = 0.$$

Taking the kinetic energy inner product of such a velocity vector with (18) gives

$$\frac{1}{2} \sum_i m_i \tilde{v}_i \cdot (\alpha \times q_i) + \frac{1}{2} \sum_i m_i \tilde{\Omega}_i \cdot I_i A_i^T \alpha = \left( \sum_i m_i q_i \times \tilde{v}_i + \sum_i A_i I_i \tilde{\Omega}_i \right) \cdot \alpha = 0$$

as required.  $\square$

Next we show that critical points and local minima of  $H(P)$  on  $\mathcal{M}_\lambda$  correspond to critical points and local minima of the amended potential  $W_\lambda(Z)$ .

**Theorem 2.**  *$P \in T\mathcal{U}$  is a critical point of  $H(P)$  on  $\mathcal{M}_\lambda$  if and only if its configuration  $Z$  is a critical point of the amended potential  $W_\lambda(Z)$  on  $\mathcal{U}$  and its velocities are the minimizing ones (18). In this case  $P$  is a local minimum of  $H$  on  $\mathcal{M}_\lambda$  if and only if  $Z$  is a local minimum of  $W_\lambda$  on  $\mathcal{U}$ . Moreover, the minimum values are equal:  $H(P) = W_\lambda(Z)$ .*

*Proof.* If  $P$  is a critical point of  $H(P)$  on  $\mathcal{M}_\lambda$  then its velocities must be a critical point of the restriction of  $H$  to  $S_\lambda(Z)$ . Since this restriction is given by a positive definite quadratic form, the only critical point is the minimum given by (18). For any  $Z \in \mathcal{U}$ , let  $P_{min}(Z) \in T\mathcal{U}$  denote the state with these minimal velocities. The energy of this state is

$$(20) \quad H(P_{min}(Z)) = W_\lambda(Z).$$

If  $P$  is a critical point of  $H$  on  $\mathcal{M}_\lambda$  then it follows that  $Z$  is a critical point of  $W_\lambda(Z)$  and if  $P$  is a local minimum of  $H$  then  $Z$  is a local minimum of  $W_\lambda$ .  $P_{min} : \mathcal{U} \rightarrow T\mathcal{U}$  will be called the *minimum energy section* of the tangent bundle.

For the converse, suppose  $Z$  is a critical point of  $W_\lambda(Z)$  in  $\mathcal{U}$  and that  $P = P_{min}(Z)$ . Then (20) shows that  $P$  is a critical point of the restriction of  $H$  to the minimal energy section and the velocities of  $P$  are critical for the restriction of  $H$  to  $S_\lambda(Z)$ . Since  $S_\lambda(Z)$  together with the tangent space to the minimal energy section span the tangent space  $T_P\mathcal{M}_\lambda$ , it follows that  $P$  is a critical point of  $H$  in  $\mathcal{M}_\lambda$ . Finally, suppose  $Z$  is a local minimum of  $W_\lambda$ . To see that  $P$  is a local minimum of  $H$  consider any curve  $P(s)$ ,  $|s| < \delta$  with  $P(0) = P$ . If  $Z(s)$  is the corresponding curve of configurations, we have

$$H(P(s)) \geq H(P_{min}(Z(s))) = W_\lambda(Z(s)) \geq W_\lambda(Z) = H(P)$$

for  $|s| < \delta$  and so  $P$  is a local minimum of  $H$  as required.  $\square$

While the amended potential appears quite naturally in the minimum energy problem, we now seek to replace it by a simpler function used by Scheeres in [6]. Recall the formula (17) for the energy  $H_\lambda(Z)$  of a relative equilibrium state in  $\mathcal{M}_\lambda$ . We will call  $H_\lambda$  the *critical energy function*. From Theorem 2 we see that  $W_\lambda(Z) = H_\lambda(Z)$  at the critical points of  $W_\lambda$ . In fact, this equation holds whenever  $e = \lambda/|\lambda|$  is an eigenvector of the total inertia matrix  $I(Z)$ . The following lemma of Scheeres [6] clarifies the relationship between the two functions.

**Lemma 3.** *For  $Z \in \mathcal{U}$  and  $\lambda \neq 0 \in \mathbb{R}^3$  we have*

$$(21) \quad H_\lambda(Z) \leq W_\lambda(Z)$$

*with equality if and only if  $\lambda$  is an eigenvector of  $I(Z)$ . Both functions provide lower bounds for the energy of any state  $P = (Z, \dot{Z}) \in \mathcal{M}_\lambda$ .*

*Proof.* We need to show that  $\lambda^T I(Z)^{-1} \lambda \geq \frac{|\lambda|^2}{G_e(Z)}$  or equivalently

$$e^T I(Z)^{-1} e \geq \frac{1}{e^T I(Z) e}$$

where  $e = \lambda/|\lambda|$  is the unit vector along  $\lambda$ . Since  $I(Z)$  is a positive definite symmetric matrix, there is a positive definite symmetric matrix  $C$  with  $I(Z) = C^2$ . Then the Cauchy-Schwarz inequality gives

$$1 = e \cdot e = (C^{-1}e) \cdot (Ce) \leq |C^{-1}e| |Ce| = (e^T I(Z)^{-1} e)^{\frac{1}{2}} \cdot (e^T I(Z) e)^{\frac{1}{2}}$$

as required. Furthermore, we have equality if and only if  $C^{-1}e$  and  $Ce$  are proportional, which means  $e$  is an eigenvector of  $C^2$ . The last statement follows from lemma 2 which also shows that the lower bound  $W_\lambda(Z)$  is sharp.  $\square$

Next we will show that  $H_\lambda$  provides an alternative variational characterization of relative equilibrium configurations.

**Theorem 3.** *The amended potential  $W_\lambda(Z)$  and the critical energy function  $H_\lambda(Z)$  have the same critical points in  $\mathcal{U}$ , namely the relative equilibrium configurations for angular momentum  $\lambda$ .*

*Proof.* Theorems 1 and 2 show that critical points of  $W_\lambda$  are exactly the relative equilibrium configurations for angular momentum  $\lambda$ . We will show that the same is true for  $H_\lambda$ . Simplify notation by writing  $G(Z)$  instead of  $G_e(Z)$ . Then the critical point equations for  $H_\lambda$  on  $\mathcal{U}$  are such that

$$-\frac{|\lambda|^2}{2G^2} G_{q_i} + m_i \beta = U_{q_i}$$

where  $\beta \in \mathbb{R}^3$  is a Lagrange multiplier, and also that

$$-\frac{|\lambda|^2}{2G^2} \frac{d}{dt} \Big|_{t=0} G(Z_i(t)) = \frac{d}{dt} \Big|_{t=0} U(Z_i(t)) \quad i = 1, \dots, n$$

where  $R(t) \in \mathbf{SO}(3)$  with  $R(0) = \mathbb{I}$  and  $Z_i(t)$  is the curve of configurations where  $A_i$  is replaced by  $A_i(t) = A_i R(t)$  and all other variables are left constant. If  $Z \in \mathcal{U}$  then summing over  $i$  in the first equation shows that  $\beta = 0$ .

Differentiating the formula (11) shows that these equations agree with the equations (15) for relative equilibrium configurations with angular momentum  $\lambda$ .  $\square$

It remains to consider the question of local minima. Assuming a certain technical condition, we will show that local minima of  $W_\lambda(Z)$  correspond to local minima of  $H_\lambda(Z)$  and vice versa. It is not clear that this condition is really necessary but we don't know how to eliminate it. We will need to use the behavior of these functions under the diagonal action of  $R \in \mathbf{SO}(3)$ . We have

$$U(RZ) = U(Z) \quad I(RZ) = RI(Z)R^T \quad I(RZ)^{-1} = RI(Z)^{-1}R^T.$$

From this we find

$$(22) \quad \begin{aligned} W_\lambda(RZ) &= W_{R^T\lambda}(Z) = \frac{1}{2}(R^T\lambda)^T I(Z)^{-1}(R^T\lambda) - U(Z) \\ H_\lambda(RZ) &= H_{R^T\lambda}(Z) = \frac{|\lambda|^2}{2G_{R^T e}(Z)} - U(Z). \end{aligned}$$

In other words, the kinetic energy terms are rotated by  $R^T$  while the potential energy term is unchanged.

Now suppose that  $Z$  is a local minimum of  $H_\lambda(Z)$ . Then the unit vector  $e$  must be a maximal eigenvector of  $I(Z)$ , that is,

$$G_e(Z) = \max_{|u|=1} u^T I(Z)u.$$

Otherwise we could find a rotation  $R$  arbitrarily close to the identity with  $G_{R^T e}(Z) > G_e(Z)$  and then (22) shows that  $Z$  is not a local minimum of  $H_\lambda$ . Similarly if  $Z$  is a local minimum of  $W_\lambda(Z)$  then  $\lambda$  must be an eigenvector of  $I(Z)$  with eigenvalue  $G_e(Z)$  which is maximal in this sense. The technical condition is that  $\pm e$  are the unique maximal eigenvectors of  $I(Z)$ , or equivalently, that the maximal eigenvalue  $G_e(Z)$  is simple.

**Lemma 4.** *Let  $Z \in \mathcal{U}$  be a configuration such that  $e$  is an eigenvector of  $I(Z)$  which is uniquely maximal in the sense that*

$$G_e(Z) = e^T I(Z)e = \max_{|u|=1} u^T I(Z)u$$

*and the maximum is achieved only at  $u = \pm e$ . Then there is a codimension-two submanifold  $\mathcal{M} \subset \mathcal{U}$  through  $Z$  such that  $e$  is a uniquely maximal eigenvector of  $I(Z')$  for all  $Z' \in \mathcal{M}$ . Moreover  $W_\lambda(Z') = H_\lambda(Z')$  for all  $Z' \in \mathcal{M}$  and there is a neighborhood  $\mathcal{V}$  of  $(Z, \mathbb{I})$  in  $\mathcal{M} \times \mathbf{SO}(3)$  such that for  $(Z', R) \in \mathcal{V}$  we have*

$$(23) \quad W_\lambda(RZ') \geq W_\lambda(Z') \quad H_\lambda(RZ') \geq H_\lambda(Z').$$

*Finally,  $Z$  is a local minimum of  $W_\lambda$  on  $\mathcal{U}$  if and only if it is a local minimum of  $H_\lambda$  on  $\mathcal{U}$ .*

*Proof.* Assume without loss of generality that  $e = (0, 0, 1)$  and that the matrix of  $I(Z)$  is diagonal:

$$I(Z) = \begin{bmatrix} I_{11} & 0 & 0 \\ 0 & I_{22} & 0 \\ 0 & 0 & I_{33} \end{bmatrix}$$

with  $I_{33} = G_e(Z)$  and  $I_{33} > \max(I_{11}, I_{22})$ . Consider the matrices  $I(Z')$  for  $Z'$  near  $Z$ . The condition that  $e$  be an eigenvector is that  $I_{13}(Z') = I_{23}(Z') = 0$ . We will use the implicit function theorem to show that these two equations define a submanifold  $\mathcal{M}$  containing  $Z$ .

Let  $R_1(t)$  be the rotation around  $(1, 0, 0)$  with unit angular speed and let  $Z'(t) = R_1(t)Z$ . Then  $I(Z'(t)) = R_1(t)I(Z)R_1(t)^T$  and we calculate

$$\frac{d}{dt}I_{13}(Z')|_{t=0} = 0 \quad \frac{d}{dt}I_{23}(Z')|_{t=0} = I_{22} - I_{33} \neq 0.$$

Similarly, if  $R_2(t)$  is the rotation around  $(0, 1, 0)$  with unit angular speed and  $Z'(t) = R_2(t)Z$ . Then

$$\frac{d}{dt}I_{13}(Z')|_{t=0} = I_{33} - I_{11} \neq 0 \quad \frac{d}{dt}I_{23}(Z')|_{t=0} = 0.$$

Note that for  $Z \in \mathcal{U}$  the rotated curves  $Z'(t)$  lie entirely in  $\mathcal{U}$ . It follows that the matrix of  $D(I_{13}, I_{23})$  on  $T_Z\mathcal{U}$  has rank 2. By the implicit function theorem, the equations  $I_{13}(Z') = I_{23}(Z') = 0$  define a local codimension-two submanifold  $\mathcal{M}$  near  $Z$ .

For each  $Z' \in \mathcal{M}$ ,  $e$  is an eigenvector of  $I(Z')$  and therefore  $W_\lambda(Z') = H_\lambda(Z')$ . By continuity,  $e$  will be uniquely maximal for  $Z'$  sufficiently close to  $Z$ . Unique maximality implies that rotating  $Z'$  will not decrease the function  $W_\lambda, H_\lambda$  so (23) holds.

Finally, suppose  $Z$  is a local minimum of one of the two functions. Since  $W_\lambda(Z') = H_\lambda(Z')$  for  $Z' \in \mathcal{M}$  both functions have local minima at  $Z$  when restricted to  $\mathcal{M}$ . The computation for the implicit function theorem shows that every point near  $Z$  in  $\mathcal{U}$  can be written as  $RZ'$  for  $(Z', R) \in \mathcal{V}$ . By (23), both functions have local minima at  $Z$ .  $\square$

Now we have most of the ingredients for our main result, namely, that for  $n \geq 3$  and  $\lambda \neq 0$ , relative equilibria are never energy minimizers in  $\mathcal{M}_\lambda$ .

**Theorem 4.** *Let  $P \in \mathcal{M}_\lambda$  be a relative equilibrium state with angular momentum  $\lambda \neq 0$ . If  $n \geq 3$  then  $P$  is not a local minimum of  $H$  on  $\mathcal{M}_\lambda$ . Equivalently, a relative equilibrium configuration  $Z \in \mathcal{U}$  is never a local minimum of the amended potential  $W_\lambda$  on  $\mathcal{U}$  for  $n \geq 3$ .*

*Proof.* By Theorem 2, it suffices to prove the statement about critical points of  $W_\lambda$  and we may assume without loss of generality that  $e = (0, 0, 1)$  and  $\lambda = (0, 0, |\lambda|) \neq 0$ .

Let  $Z \in \mathcal{U}$  be a relative equilibrium configuration for angular momentum  $\lambda$ . First consider the case where  $Z$  satisfies the technical condition of lemma 4. Then it suffices to show that  $Z$  is not a local minimum of the simpler function  $H_\lambda$ .

Let  $Z \in \mathcal{U}$  be any critical point of  $H_\lambda$ . We will construct a curve in configuration space  $Z(s) \in \mathcal{U}$ ,  $|s| < \delta$  with  $Z(0) = Z$  as follows. We will leave the orientation matrices  $A_i$  constant and the positions of the centers of mass will have the form  $q_i(s) = q_i + sw_i$  for some vectors  $w_i \in \mathbb{R}^3$  with  $m_1w_1 + \dots + m_nw_n = 0$ . The vector  $w = (w_1, \dots, w_n) \in \mathbb{R}^{3n}$  will be chosen such that  $D_q^2H_\lambda(w, w) < 0$  where  $D_q^2H_\lambda$  is the Hessian of  $H_\lambda$  with respect to the  $q_i$  variables. Since  $Z = Z(0)$  is a critical point of  $H_\lambda$  we have

$$D_qH_\lambda(Z)w = 0 \quad D_q^2H_\lambda(w, w) < 0.$$

For  $\delta > 0$  sufficiently small, we will have  $H_\lambda(Z(s)) < H_\lambda(Z(0))$  for  $|s| < \delta, s \neq 0$ , showing that  $Z$  is not a local minimum.

For simplicity we will write  $G(Z)$  instead of  $G_e(Z)$ . We have

$$DH_\lambda(Z)w = -\frac{|\lambda|^2}{2G(Z)^2}G_q(Z) \cdot w - U_q(Z) \cdot w = 0$$

and

$$(24) \quad \begin{aligned} D_q^2 H_\lambda(w, w) &= \frac{|\lambda|^2}{G(Z)^3} (G_q(Z) \cdot w)^2 \\ &\quad - \frac{|\lambda|^2}{2G(Z)^2} D_q^2 G(Z)(w, w) - D_q^2 U(Z)(w, w). \end{aligned}$$

Let  $M$  be the  $3n \times 3n$  block diagonal matrix with  $3 \times 3$  blocks  $m_i \mathbb{I}$ . The vector  $w$  will be an eigenvector of the matrix  $B = M^{-1} D_q^2 H_\lambda$  with a negative eigenvalue  $\mu < 0$ . The three terms in (24) give a decomposition  $B = B_1 + B_2 + B_3$ . Since

$$G_q(Z) = 2(m_1 K_e q_1, \dots, m_n K_e q_n)$$

the matrix  $B_1$  breaks up into  $3 \times 3$  blocks  $b_{ij}^1$  where

$$b_{ij}^1 = \frac{4m_j |\lambda|^2}{G(Z)^3} \begin{bmatrix} x_i x_j & x_i y_j & 0 \\ y_i x_j & y_i y_j & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The matrix  $B_2$  is block diagonal with  $3 \times 3$  diagonal blocks

$$b_{ii}^2 = -\frac{|\lambda|^2}{G(Z)^2} K_e = -\frac{|\lambda|^2}{G(Z)^2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Finally, using (4) we find the third term  $B_3 = -M^{-1} D_q^2 U(Z)$  breaks up into  $3 \times 3$  blocks

$$b_{ij}^3 = -\frac{1}{m_i} \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{dm_i dm_j}{r_{ij}^3} (-\mathbb{I} + 3u_{ij} u_{ij}^T) \quad i \neq j$$

and diagonal blocks

$$b_{ii}^3 = -\sum_{j \neq i} b_{ij}^3.$$

All of the blocks of  $B_3$  have zero trace, essentially due to the fact that the Newtonian potential is a harmonic function on  $\mathbb{R}^3$ . Calculating the traces of  $B_1, B_2$  we get

$$\text{trace}(B) = \frac{(4\theta - 2n)|\lambda|^2}{G(Z)^2}$$

where

$$\theta = \frac{\sum_j m_j (x_j^2 + y_j^2)}{G(Z)}.$$

Now the formula (11) for  $G(Z)$  includes the sum in the numerator plus other positive terms. It follows that  $0 \leq \theta < 1$  and therefore

$$\text{trace}(B) < \frac{(4 - 2n)|\lambda|^2}{G(Z)^2}$$

The mass matrix  $M$  defines an inner product on  $\mathbb{R}^{3n}$ :

$$\langle v, w \rangle = v^T M w.$$

$B$  is an  $M$ -symmetric matrix so its eigenvalues are all real and its eigenvectors are orthogonal with respect to this inner product. Let  $\hat{e}_1 = (e_1, e_1, \dots)$  where  $e_1 = (1, 0, 0)$  and define  $\hat{e}_2, \hat{e}_3$  similarly. An easy computation shows that  $\hat{e}_i$  are eigenvectors of  $B$  with eigenvalues

$$\mu_1 = \mu_2 = -\frac{|\lambda|^2}{G(Z)^2} \quad \mu_3 = 0.$$

Note that the  $M$ -orthogonal complement of the span of the  $\hat{e}_i$  is exactly the zero center of mass subspace and it is an invariant subspace for  $B$ . Let  $\mu_4, \dots, \mu_{3n}$  be the eigenvalues of  $B$  on this subspace. Then we have

$$\mu_4 + \dots + \mu_{3n} < \frac{(6 - 2n)|\lambda|^2}{G(Z)^2}.$$

Since  $n \geq 3$  this sum is strictly less than zero and we have a negative eigenvalue, as required.

To finish, we need to rule out the possibility of local minima for which  $e$  is not uniquely maximal. This time we have to work directly with the amended potential  $W_\lambda$ . If  $e$  is not uniquely maximal then without loss of generality we may assume that the total inertia tensor takes one of the two forms:

$$I(Z) = \begin{bmatrix} I_{11} & 0 & 0 \\ 0 & I_{33} & 0 \\ 0 & 0 & I_{33} \end{bmatrix} \quad I(Z) = \begin{bmatrix} I_{33} & 0 & 0 \\ 0 & I_{33} & 0 \\ 0 & 0 & I_{33} \end{bmatrix} = I_{33} \mathbb{I}.$$

We will show that these conditions together with the relative equilibrium equations (15) put strong restrictions on the configuration.

First suppose  $Z$  is a local minimum of  $W_\lambda$  with  $I(Z) = I_{33} \mathbb{I}$ . Then (22) shows that  $W_\lambda(RZ) = W_\lambda(Z)$  for all  $R \in \mathbf{SO}(3)$ . Since  $Z$  is a local minimum, the rotated configurations  $RZ$  with  $R$  sufficiently close to  $\mathbb{I}$  must also be local minima. In particular  $RZ$  is a critical point of  $W_\lambda$ . By rotational symmetry,  $Z = R^T(RZ)$  must a critical point of  $W_{R^T \lambda}$  in addition to being a critical point of  $W_\lambda$ . Now the first equation of (15) shows that  $U_{q_i} \in e^\perp$  and the corresponding equation for  $R^T \lambda$  shows that  $U_{q_i} \in (R^T e)^\perp$  for all  $R \in \mathbf{SO}(3)$  sufficiently close to  $\mathbb{I}$ . But this implies  $U_{q_i} = 0$ . Therefore the projections of the position vectors  $q_i$  onto all of the  $(R^T e)^\perp$  must vanish and therefore  $q_i = 0$  for all  $i = 1, \dots, n$ . Clearly it is a very special type of relative equilibrium where the bodies are disjoint, but they all have the same center of mass. An example would be nested spherical shells of mass.

In the case where  $I_{33} = I_{22} > I_{11}$ , a similar argument applies using rotations  $R$  around  $(1, 0, 0)$ . The conclusion is that  $Z$  must be a relative equilibrium not just for  $\lambda$  but also for  $R^T\lambda$  and that the projections of the  $q_i$  onto all of the subspaces  $(R^T e)^\perp$  must vanish. In this case, all of the centers of mass  $q_i$  are collinear and lie on the first coordinate axis. In other words,  $q_i = (x_i, 0, 0)$ . The previous case can be subsumed into this one by taking  $x_i = 0$ .

To show that local minima are impossible in these two cases, consider the Hessian  $D_q^2 W_\lambda$  of  $W_\lambda$  with respect to  $q$ . As before, our goal will be to find a vector  $w$  in the zero momentum subspace such that  $D_q^2 W_\lambda(w, w) < 0$ . After some computation we find that the formula for  $D_q^2 W_\lambda$  agrees with formula (24) for  $D_q^2 H_\lambda$  except that the first term is replaced by the more complicated expression

$$\frac{|\lambda|^2}{2G(Z)^2} (D_q I(Z)(w)e)^T I(Z)^{-1} (D_q I(Z)(w)e).$$

This term is positive semi-definite. We will eliminate it by choosing a vector  $w$  in the subspace such that  $D_q I(Z)(w)e = 0$ . However, to make the rest of the proof work, it will be important to use subspaces which are invariant under the diagonal action of the rotation group  $\mathbf{SO}(3)$ . To this end, we also require  $D_q I(Z)(Rw)e = 0$  for every rotation  $R \in \mathbf{SO}(3)$ .

Differentiating (10) with respect to the  $q_i$  at  $q_i = (x_i, 0, 0)$  and recalling that  $e = (0, 0, 1)$ , we find

$$D_q I(Z)(w)e = \sum_i m_i \begin{bmatrix} 0 & 0 & -x_i \\ 0 & 0 & 0 \\ 2x_i & 0 & 0 \end{bmatrix} \begin{bmatrix} w_{i1} \\ w_{i2} \\ w_{i3} \end{bmatrix}.$$

Define vectors  $v_1 = (0, 0, -x_1, 0, 0, -x_2, \dots)$  and  $v_2 = (x_1, 0, 0, x_2, 0, 0, \dots)$  in  $\mathbb{R}^{3n}$ . To complete the proof we want to restrict  $w$  to a rotation invariant subspace such that  $D_q I(Z)(w)e$ . We can use  $\mathcal{G}^\perp$  where

$$\mathcal{G} = \text{span}\{Rv_1, Rv_2 : R \in \mathbf{SO}(3)\} \subset \mathbb{R}^{3n}.$$

Note that  $v_1$  and  $v_2$  are actually in the same orbit of the diagonal action of  $\mathbf{SO}(3)$  so we can use just one of them, say  $v_2$  in the definition of  $\mathcal{G}$ . Note that all of the vectors  $(x_i, 0, 0) \in \mathbb{R}^3$  can be expressed as linear combination of just one of them. These dependence relations define a three-dimensional subspace of  $\mathbb{R}^{3n}$ . The rotated vectors  $Rv_2$  satisfy the same dependence relations. Therefore  $\mathcal{G}$  is contained in this three-dimensional subspace and  $\dim \mathcal{G} \leq 3$ . Note that the zero momentum subspace is also rotation invariant and contains  $\mathcal{G}$ . Taking the orthogonal complement of  $\mathcal{G}$  within the zero momentum space gives a rotation invariant subspace of dimension  $\dim \mathcal{G}^\perp \geq 3n - 6 \geq 3$ .

Choose any nonzero vector  $w \in \mathcal{G}^\perp$  and consider the average of the quadratic form  $D_q^2 W_\lambda(Z)(Rw, Rw)$  as  $R$  runs over the rotation group  $\mathbf{SO}(3)$ .

Then using the last two terms of (24) and (4) we have

$$D_q^2 W_\lambda(w, w) = -\frac{|\lambda|^2}{G(Z)^2} \sum_i m_i w_i^T K_e w_i - \sum_{i < j} \int_{\mathcal{B}_i} \int_{\mathcal{B}_j} \frac{dm_i dm_j}{r_{ij}^3} (-|w_{ij}|^2 + 3(u_{ij} \cdot w_{ij})^2).$$

This expresses  $D_q^2 W_\lambda(w, w)$  as a sum of quadratic form on  $\mathbb{R}^3$ . Since the rotation group acts diagonally, we can find the average of  $D_q^2 W_\lambda(Z)(Rw, Rw)$  as a sum of averages of these three-dimensional quadratic forms. Since  $\mathbf{SO}(3)$  acts orthogonally and irreducibly on  $\mathbb{R}^3$ , it follows from Schur's lemma that these averaged forms are just scalar multiples of the identity where the scalar is the trace of the matrix representing the form.

Since  $K_e$  is orthogonal projection onto a plane, its average is  $2\mathbb{I}$ . The quadratic forms in  $w_{ij}$  are given by integrals but their traces are all zero. Hence the average over  $R \in \mathbf{SO}(3)$  of the forms  $D_q^2 W_\lambda(Rw, Rw)$  is

$$\overline{D_q^2 W_\lambda} = -\frac{|\lambda|^2}{G(Z)^2} \sum_i 2m_i |w_i|^2 < 0.$$

Hence there is a vector of the form  $w' = Rw$  with  $D_q^2 W_\lambda(w', w') < 0$  and the proof is complete. □

## REFERENCES

1. V.I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer-Verlag, New York, (1989).
2. A.J. Maciejewski, *Reduction, relative equilibrium and potential in the two rigid bodies problem*, Cel.Mech.Dyn.Ast. 63(1) (1995) 1–28.
3. W.D. MacMillan, *Theory of the Potential*, Dover, New York, (1958).
4. J. Marsden, *Lecture on Mechanics*, London Mathematical Society Lecture Notes Series **174**, Cambridge University Press (1992).
5. R. Moeckel, *On central configurations*, Math. Zeit. **205** (1990) 499–517.
6. D.J. Scheeres, *Minimum energy configurations in the N-body problem and the celestial mechanics of granular systems*, Cel.Mech.Dyn.Ast. 113 (2012) 291–320.
7. J.C. Simo, D. Lewis, J.E. Marsden, *Stability of relative equilibria. Part i: The reduced energy-momentum method*, Arch.Rat.Mech. 115(1)(1991) 15–59.
8. S. Smale, *Topology and mechanics, I*, Inv. Math., **10**, (1970) 305–331.
9. S. Smale, *Topology and Mechanics, II, The planar n-body problem*, Invent. Math. **11** (1970) 45–64.

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS MN 55455  
E-mail address: rick@math.umn.edu