Two rolling disks or spheres

George W. Patrick and Tyler Helmuth

Applied Mathematics and Mathematical Physics
Department of Mathematics and Statistics
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

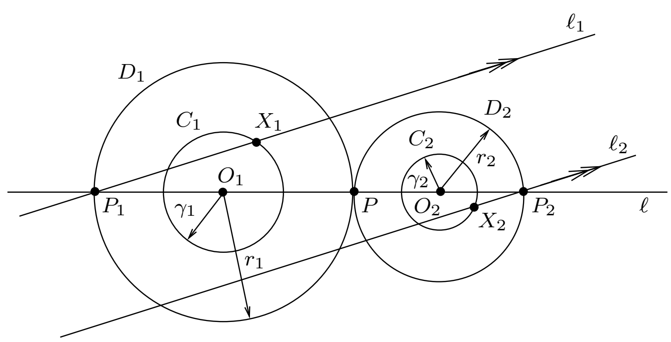
July 31, 2009

Abstract

The mechanical system of two disks, moving freely in the plane, while in contact and rolling against each other without slipping, may be written as a Lagrangian system with three degrees of freedom and one holonomic rolling constraint. We derive simple geometric criteria for the rotational relative equilibria and their stability. Extending to three dimensions, we derive the kinematics of the analogous system where two spheres replace two disks, and we verify that the rolling disk system occurs as a holonomic subsystem of the rolling sphere system.

1 Rolling disks

In the figure below, C_k and D_k are concentric (at O_k) circles of radius γ_k and r_k respectively, $k = 1, 2$. D_1 and D_2 are tangent at P, so that O_1, O_2, P are colinear and lie on the line ℓ.



Let point P_k be the intersection of ℓ and D_k diametrically opposite to P. We prove the following pretty fact: two disks, rolling against each other in the plane, with centers of mass at X_1 and X_2, are in rotational relative equilibrium if and only if the lines ℓ_1 and ℓ_2 are parallel, as shown. Moreover, with the rolling constraint imposed, if the ordered pair $X_1 - P_1, X_2 - P_2$ changes from a negative orientation to a positive orientation as disk $k = 1$ is rotated counterclockwise, then the relative equilibrium is stable. If the opposite condition is true, then the relative equilibrium is unstable. Stable relative equilibria are motions such that the entire assemblage is balanced and rotates stably and freely, with no control, as a rigid unit.

See [3, 7] for general information and references concerning nonholonomic mechanics in the geometric spirit of this article. Define $\mathrm{SO}(2) = \big\{ z \in \mathbb{C} : |z| = 1 \big\}$. The system admits an elegant coordinate free Lagrangian formulation, where the configuration of disk k is parametrized, using elements of $\mathrm{SO}(2) \times \mathbb{C}$ and a reference disk in the plane $\mathbb{C}$; by definition, in the configuration (z_k, a_k), a point p in disk k has location $z_k p + a_k$. Let reference disk k have center at the origin, radius r_k, mass m_k, center of mass at γ_k, and moment of inertia I_k. Let $u_k \in \mathrm{SO}(2)$ be such that the reference configuration contact point of the two disks is at $r_k u_k$. Not every element of $(\mathrm{SO}(2) \times \mathbb{C})^2$ represents a physical configuration of this system, since the disks must touch each other at the contact point in such a way that there is no interpenetration. This gives the constraints

$$r_1 z_1 u_1 + a_1 = r_2 z_2 u_2 + a_2, \qquad z_1 u_1 = -z_2 u_2. \tag{1}$$

Using the variable $u \in \mathrm{SO}(2)$ defined by $u \equiv z_1 u_1 \equiv -z_2 u_2$ reduces the two variables u_1 and u_2 to one, with

$$u_1 = z_1^{-1} u, \qquad u_2 = -z_2^{-1} u. \tag{2}$$

Also, assume the system is in the frame where the center of mass is at the origin i.e. impose that the configurations (z_k, a_k) satisfy

$$m_1(z_1\gamma_1 + a_1) + m_2(z_2\gamma_1 + a_2) = 0. \tag{3}$$

Solving Equations (1) and (3) for a_1 and a_2 gives

$$\begin{aligned} a_1 &= -\frac{1}{m_1 + m_2}\big(m_1 z_1 \gamma_1 + m_2 z_2 \gamma_2 + R m_2 u\big), \\ a_2 &= -\frac{1}{m_1 + m_2}\big(m_1 z_1 \gamma_1 + m_2 z_2 \gamma_2 - R m_1 u\big), \end{aligned} \tag{4}$$

where $R \equiv r_1 + r_2$. Thus the system has configuration space the three torus $\mathcal{Q} \equiv \mathrm{SO}(2)^3 = \big\{ (z_1, z_2, u) \big\}$ i.e. there are three degrees of freedom.

Required also is the no-slip rolling constraint i.e. the condition that the instantaneous velocities of the points on the disks corresponding to the contact points u_k are the same. In equations, this is

$$\frac{d}{dt}\bigg|_{u_1,\, u_2 \text{ constant}} \Big(r_1 z_1 u_1 + a_1 = r_2 z_2 u_2 + a_2\Big). \tag{5}$$

Subtracting from this the derivative of Equation (1), where u_1 and u_2 are not constant, gives $r_1 z_1 \dot{u}_1 = r_2 z_2 \dot{u}_2$, which is view of (2) is

$$R u^{-1} \dot{u} - r_1 z_1^{-1} \dot{z}_1 - r_2 z_2^{-1} \dot{z}_2 = 0.$$

Evidently, the rolling constraint for this system is actually holonomic, because it (locally) integrates to the SO(2) valued function $u^R/z_1^{r_1}z_2^{r_2}$.

The Lagrangian for the system is the sum of the kinetic energies of the individual disks. If the mass density of disk k is $\rho_k(p)$, then this is

$$\begin{aligned} KE_k &= \frac{1}{2}\int d\rho_k(p)|\dot{z}_k p + \dot{a}_k|^2 \\ &= \frac{1}{2}\left(\int d\rho_k(p)|p|^2\right)|\dot{z}_k|^2 + \mathrm{Re}\left(\left(\int d\rho_k(p)p\right)\dot{z}_k\dot{a}_k^-\right) \\ &\qquad\qquad + \frac{1}{2}\int d\rho_k(p)|\dot{a}_k|^2 \\ &= \frac{I_k}{2}|\dot{z}_k|^2 + m_k\,\mathrm{Re}(\gamma_k\dot{z}_k\dot{a}_k^-) + \frac{m_k}{2}|\dot{a}_k|^2. \end{aligned} \tag{6}$$

The group SO(2) acts on the variables z_k, a_k, u by complex multiplication. The action on the variables $\dot{z}_k, \dot{a}_k, \dot{u}$ is also multiplication, and the kinetic energy of each disk is invariant. So the Lagrangian is invariant and the system admits this SO(2) symmetry. The action on the variables u_k is trivial.

Since the system is holonomic, the relative equilibria may be computed using the amended potential, which is the Lagrangian evaluated on the infinitesimal generator of the action [6]. The infinitesimal generator of $\xi \in \mathbb{R}$ is given by $\dot{z}_k = i\xi z_k$ and similarly $\dot{a}_k = i\xi a_k$. Substitution gives the amended potential

$$\begin{aligned} V_\xi &= \xi^2\,\mathrm{Re}(m_1\gamma_1 z_1 a_1^- + m_2\gamma_2 z_2 a_2^-) + \frac{m_1}{2}\xi^2|a_1|^2 + \frac{m_2}{2}\xi^2|a_2|^2 + \text{constant} \\ &= -\frac{\xi^2 m_1 m_2}{m_1+m_2}\,\mathrm{Re}\big(\gamma_1\gamma_2^- z_1 z_2^- + \gamma_1 R z_1 u^- \\ &\qquad\qquad + \gamma_2\gamma_1^- z_2 z_1^- - \gamma_2 R z_2 u^- - \gamma_1 z_1(\gamma_2 z_2)^-\big) + \text{constant} \\ &= -\frac{\xi^2 m_1 m_2}{m_1+m_2}\,\mathrm{Re}\big(\gamma_1\gamma_2^- z_1 z_2^- + R(\gamma_1 z_1 - \gamma_2 z_2)u^-\big) + \text{constant}. \end{aligned}$$

By rotational symmetry one may assume $u = 1$. A necessary and sufficient condition for a relative equilibrium is the vanishing of the derivative in the direction of the rolling constraint i.e. in the direction of $r_2 - r_1 i$. The derivative of V_ξ, with respect to z_k, may be computed as the derivative with respect to θ at $\theta = 0$ of the expression obtained by replacing z_k in V_ξ with $e^{i\theta}z_k$. Assuming $\xi \neq 0$, deleting the positive constant multiplier $\xi^2/(m_1+m_2)$, and tracing the derivative through Equations (4), gives the relative equilibria as the solutions to the equations

$$\begin{aligned} 0 &= \mathrm{Im}\big((\gamma_1\gamma_2^- z_1 z_2^- + R\gamma_1 z_1)r_2 + (\gamma_1\gamma_2^- z_1 z_2^- + R\gamma_2 z_2)r_1\big) \\ &= R\,\mathrm{Im}\big((\gamma_1\gamma_2^- z_1 z_2^- + r_2\gamma_1 z_1 + r_1\gamma_2 z_2\big) \\ &\quad = R\,\mathrm{Im}\big((z_1\gamma_1 + r_1)(z_2\gamma_2 - r_2)^-\big). \end{aligned}$$

This is as required because, referring to the diagram at the beginning,

$$\begin{aligned} X_1 - P_1 &= (X_1 - O_1) - (O_1 - P_1) = z_1\gamma_1 + r_1, \\ X_2 - P_2 &= (X_2 - O_2) - (O_2 - P_2) = z_2\gamma_2 - r_2, \end{aligned}$$

and lines ℓ_k are along $X_k - P_k$ because, generally, $z, w \in \mathbb{C}$ are linearly dependent if and only if $\mathrm{Im}(zw^-) = 0$. Also, the derivative of V_ξ along the rolling constraint in the direction of counterclockwise rotation of disk $k = 1$ is a positive constant of $\mathrm{Im}\big((X_1 - P_1)(X_2 - P_2)^-\big)$, and the relative equilibria are stable at local minima of the amended potential i.e. when this quantity changes sign from negative to positive.

2 Rolling spheres

A natural generalization is the system where two spheres roll freely against one another in three space. This is a nonholonomic version of the system where two rigid bodies are coupled with an ideal ball-and-socket joint [5].

Consider two three-dimensional reference spheres with, as before, centers at the origin, radii r_k, masses m_k, and centers of mass at γ_k, $k = 1, 2$. Parametrize the configurations of sphere k by $(z_k, a_k) \in \mathrm{SO}(3) \times \mathbb{R}^3$. Let $u_k \in S^2$ be such that the reference configuration contact point of the two disks is at $r_k u_k$. Equations (1)–(4) are unchanged, and, after elimination of u_k, the configuration space is $\mathcal{Q} \equiv \mathrm{SO}(3)^2 \times S^2 = \big\{(z_1, z_2, u)\big\}$.

If $c \in \mathbb{R}^3$, then let $c^\wedge$ be the 3×3 matrix which represents the linear map $x \mapsto c \times x$, and identify $\boldsymbol{T}\mathrm{SO}(3) = \mathrm{SO}(3) \times \mathbb{R}^3 = \big\{(z, \Omega)\big\}$ by left translation i.e. (z, Ω) is the tangent vector corresponding to the derivative at $\epsilon = 0$ of the curve $z \exp(\Omega^\wedge \epsilon)$, and

$$\boldsymbol{T}\mathcal{Q} = \Big\{ \big((z_1, z_2, u), (\Omega_1, \Omega_2, \dot{u})\big) : (z_1, z_2, u) \in \mathcal{Q},\ \Omega_1, \Omega_2, \dot{u} \in \mathbb{R}^3,\ u \cdot \dot{u} = 0 \Big\}.$$

The no slip rolling constraint is $r_1 z_1 \dot{u}_1 = r_2 z_2 \dot{u}_2$ where $\dot{u}_k$ are determined from (2) and the curve representing the tangent vector. Since $u_1 = z_1^{-1} u$,

$$\begin{aligned} r_1 z_1 \dot{u}_1 &= r_1 z_1 \frac{d}{dt}\bigg|_{t=0} \big(z_1 \exp(\Omega_1^\wedge t)\big)^{-1} u \\ &= -r_1 z_1 \Omega_1^\wedge z_1^{-1} u + r_1 z_1 z_1^{-1} \dot{u} = -r_1 (z_1 \Omega_1)^\wedge u + r_1 \dot{u}, \end{aligned}$$

and similarly $r_2 z_2 \dot{u}_2 = r_2 (z_2 \Omega_2)^\wedge u - r_2 \dot{u}$, the rolling constraint is

$$r_1 z_1 \dot{u}_1 - r_2 z_2 \dot{u}_2 = R\dot{u} - (r_1 z_1 \Omega_1 + r_2 z_2 \Omega_2) \times u = 0. \tag{7}$$

The mass and center of mass are

$$m_k \equiv \int \rho_k(p)\, dp, \qquad \gamma_k \equiv \frac{1}{m_k} \int p \rho_k(p)\, dp.$$

Let I_k be the coefficient of inertia matrix of body i in the reference configuration with respect to the center of mass:

$$\begin{aligned} I_k &= \int (p - \gamma_k)(p - \gamma_k)^t \rho_k(p)\, dp \\ &= \int \left[p p^t - \gamma_k p^t - p {\gamma_k}^t + \gamma_k {\gamma_k}^t \right] \rho_k(p)\, dp = \int p p^t \rho_k(p)\, dp - m_k \gamma_k \gamma_k^t. \end{aligned}$$

The kinetic energy of sphere k is

$$\begin{aligned}
KE_k &\equiv \frac{1}{2}\int |\dot{z}_k p + \dot{a}_k|^2 \rho_k(p)\,dp \\
&= \frac{1}{2}\int \left(|\dot{z}_k p|^2 + 2\dot{z}_k p\cdot\dot{a}_k + |\dot{a}_k|^2\right)\rho_k(p)\,dp \\
&= \frac{1}{2}\int \left(\mathrm{trace}\big((\dot{z}_k p)(\dot{z}_k p)^t\big) + 2\dot{z}_k p\cdot\dot{a}_k + |\dot{a}_k|^2\right)\rho_k(p)\,dp \\
&= \frac{1}{2}\,\mathrm{trace}\left(\dot{z}_k\left[\int pp^t\rho_k(p)\,dp\right]\dot{z}_k{}^t\right) + \dot{z}_k\left[\int p\rho_k(p)\,dp\right]\cdot\dot{a}_k + \frac{m_k}{2}|\dot{a}_k|^2 \\
&= \frac{1}{2}\,\mathrm{trace}(\dot{z}_k I_k \dot{z}_k^t) + \frac{m_k}{2}\,\mathrm{trace}(\dot{z}_k\gamma_k\gamma_k^t\dot{z}_k^t) + m_k(\dot{z}_k\gamma_k)\cdot\dot{a}_k + \frac{m_k}{2}|\dot{a}_k|^2 \\
&= \frac{1}{2}\,\mathrm{trace}(\dot{z}_k I_k \dot{z}_k^t) + \frac{m_k}{2}|\dot{z}_k\gamma_k + \dot{a}_k|^2
\end{aligned}$$

i.e. the kinetic energy of rigid rotation plus the kinetic energy of the motion of the center of mass. Using (4),

$$\begin{aligned}
\dot{z}_1\gamma_1 + \dot{a}_1 &= \frac{m_2}{m_1+m_2}\big(\dot{z}_1\gamma_1 - \dot{z}_2\gamma_2 - R\dot{u}\big), \\
\dot{z}_2\gamma_2 + \dot{a}_2 &= -\frac{m_1}{m_1+m_2}\big(\dot{z}_1\gamma_1 - \dot{z}_2\gamma_2 - R\dot{u}\big),
\end{aligned}$$

and also $\mathrm{trace}(\dot{z}_k I_k \dot{z}_k^t) = \Omega^t J_k \Omega$ where $J_k^0 \equiv (\mathrm{trace}\, I_k)\mathbf{1} - I_k$ are the *moment of inertia tensors*. So, defining

$$e \equiv \frac{m_1 m_2}{m_1+m_2}, \quad J_k \equiv \frac{1}{eR^2}\big(J_k^0 + e(|\gamma_k|^2\mathbf{1} - \gamma_k\gamma_k^t)\big), \quad e_k \equiv \frac{1}{r_k}\gamma_k, \quad \alpha_k = \frac{r_k}{R}$$

the kinetic energy is,

$$\begin{aligned}
KE &= \frac{1}{2}\Omega_1^t J_1^0\Omega_1 + \frac{1}{2}\Omega_2^t J_1^0\Omega_2 + \frac{e}{2}\big|R\dot{u} - z_1(\Omega_1\times\gamma_1) + z_2(\Omega_2\times\gamma_2)\big|^2 \\
&= \frac{1}{2}\Omega_1^t J_1^0\Omega_1 + \frac{1}{2}\Omega_2^t J_1^0\Omega_2 + \frac{e}{2}|\Omega_1\times\gamma_1|^2 + \frac{e}{2}|\Omega_2\times\gamma_2|^2 + \frac{eR^2}{2}|\dot{u}|^2 \\
&\qquad - eR\dot{u}\cdot\big(z_1(\Omega_1\times\gamma_1) - z_2(\Omega_2\times\gamma_2)\big) - ez_1(\Omega_1\times\gamma_1)\cdot z_2(\Omega_2\times\gamma_2) \\
&= \frac{eR^2}{2}\Omega_1^t J_1\Omega_1 + \frac{eR^2}{2}\Omega_2^t J_1\Omega_2 + \frac{eR^2}{2}|\dot{u}|^2 \\
&\qquad - eR^2\dot{u}\cdot\big(\alpha_1 z_1(\Omega_1\times e_1) - \alpha_2 z_2(\Omega_2\times e_2)\big) \\
&\qquad\quad - eR^2\alpha_1\alpha_2\, z_1(\Omega_1\times e_1)\cdot z_2(\Omega_2\times e_2).
\end{aligned}$$

Rescaling, one can use for this system the Lagrangian

$$\begin{aligned}
L \equiv\ & \frac{1}{2}\Omega_1^t J_1\Omega_1 + \frac{1}{2}\Omega_2^t J_2\Omega_2 + \frac{1}{2}|\dot{u}|^2 \\
& - \dot{u}\cdot\big(\alpha_1 z_1(\Omega_1\times e_1) - \alpha_2 z_2(\Omega_2\times e_2)\big) \\
& \quad - \alpha_1\alpha_2\, z_1(\Omega_1\times e_1)\cdot z_2(\Omega_2\times e_2).
\end{aligned} \tag{8}$$

2.1 Symmetry

The group SO(3) acts on $\mathcal{Q}$ tridiagonally by

$$A \cdot (z_1, z_2, u) = (Az_1, Az_2, Au)$$

and this action lifts to

$$A \cdot (z_1, z_2, u, \Omega_1, \Omega_2, \dot{u}) = (Az_1, Az_2, u, \Omega_1, \Omega_2, A\dot{u}),$$

under which L is invariant. The left translated infinitesimal generator of Ω for the action on $\mathcal{Q}$ is

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \big(z_1^{-1}\exp(\Omega^\wedge\epsilon)z_1, z_2^{-1}\exp(\Omega^\wedge\epsilon)z_2, \exp(\Omega^\wedge\epsilon)u\big) = (z_1^{-1}\Omega, z_2^{-1}\Omega, \Omega \times u).$$

This satisfies the rolling constraint because

$$\big(\alpha_1 z_1(z_1^{-1}\Omega) + \alpha_2 z_2(z_2^{-1}\Omega)\big) \times u = (\alpha_1 + \alpha_2)\Omega \times u = \Omega \times u.$$

Thus the symmetry is horizontal and the Noether momentum associated to SO(3)—total angular momentum—is conserved [2, 7].

If e_k is an eigenvector of J_k, $k = 1, 2$, then the spheres have the material symmetries $\big\{(A_1, A_2) : A_k e_k = e_k\big\} \cong \mathrm{SO}(2)^2$ of rotation about the axes e_k, corresponding to the right action

$$(A_1, A_2) \cdot (z_1, z_2, u) = (z_1 A_1^{-1}, z_2 A_2^{-1}, u),$$

with lift

$$(A_1, A_2) \cdot (z_1, z_2, u, \Omega_1, \Omega_2, \dot{u}) = (z_1 A_1^{-1}, z_2 A_2^{-1}, u, A_1\Omega_1, A_2\Omega_2, \dot{u}).$$

The Lagrangian is invariant under this symmetry because, for example

$$\begin{aligned}
&(A_1\Omega_1)^t J_1 A_1 \Omega_1 = \Omega_1^t A_1^t J_1 A_1 \Omega_1 = \Omega_1^t J_1 \Omega_1, \\
&z_1 A_1^{-1}(A_1\Omega_1 \times e_1) = z_1(\Omega_1 \times A_1^{-1} e_1) = z_1(\Omega_1 \times e_1).
\end{aligned}$$

However, the infinitesimal generator of (σ_1, σ_2) is $(\sigma_1, \sigma_2, 0)$, which does not satisfy the rolling constraint, so the associated Noether momentum is not conserved.

2.2 Curvature

We will require a formula for the Lie bracket of vector field on the product of manifolds. Let M_1, M_2 be manifolds and let $X = (X^1, X^2)$, $\tilde{X} = (\tilde{X}^1, \tilde{X}^2)$ be two vector fields on $M_1 \times M_2$. Then

$$\begin{aligned}
[X, \tilde{X}] &= [(X^1, 0) + (0, X^2), (\tilde{X}^1, 0) + (0, \tilde{X}^2)] \\
&= ([X^1, \tilde{X}^1], 0) + (0, [X^2, \tilde{X}^2]) + [(X^1, 0), (0, \tilde{X}^2)] + [(0, X^2), (\tilde{X}^1, 0)].
\end{aligned}$$

The first two terms are the Lie brackets on $\mathcal{M}_k$ with the other variable held fixed. For the third term, let g be a function only of $x^1 \in M_1$ and compute

$$\begin{aligned}\big([(X^1,0),(0,\tilde{X}^2)]\,g\big)(x_0^1,x_0^2) &= \big((X^1,0)(0,\tilde{X}^2)\,g - (0,\tilde{X}^2)(X^1,0)\,g\big)(x_0^1,x_0^2)\\ &= -\big(0,\tilde{X}^2\big)\big((x^1,x^2)\mapsto \boldsymbol{d}g(x_0^1)\,X^1\big)(x_0^1,x^2)\\ &= -\boldsymbol{d}g(x_0^1)\frac{\partial X^1}{\partial x_2}\tilde{X}^2(x_0^1,x_0^2),\end{aligned}$$

where

$$\frac{\partial X^1}{\partial x_2}\tilde{X}^2(x_0^1,x_0^2) \equiv \frac{d}{dt}\bigg|_{t=0} X^1\big(x_0^1, x_0^2 + t\tilde{X}^2(x_0^1,x_0^2)\big)$$

($(x_0^1,x_0^2)+t\tilde{X}^2(x_0^1,x_0^2)$ means any curve tangent to $\tilde{X}^2(x_0^1,x_0^2)$ at $t=0$ and the derivative is the usual limit of the difference because the curve occurs in a single tangent space). Similarly, if g is a function of only x^2, then

$$\big([(X^1,0),(0,\tilde{X}^2)]\,g\big)(x_0^1,x_0^2) = \boldsymbol{d}g(x_0^2)\frac{\partial \tilde{X}^2}{\partial x_1}X^1(x_0^1,x_0^2)$$

with the result that

$$\big([(X^1,0),(0,\tilde{X}^2)] = \left(-\frac{\partial X^1}{\partial x_2}\tilde{X}^2, \frac{\partial \tilde{X}^2}{\partial x_1}X^1\right).$$

From this,

$$[(0,X^2),(\tilde{X}^1,0)] = -[(\tilde{X}^1,0),(0,X^2)] = \left(\frac{\partial \tilde{X}^1}{\partial x_2}X^2, -\frac{\partial X^2}{\partial x_1}\tilde{X}^1\right)$$

so

$$\begin{aligned}[X,\tilde{X}] = \bigg([X^1,\tilde{X}^1] + \frac{\partial \tilde{X}^1}{\partial x_2}X^2 - \frac{\partial X^1}{\partial x_2}\tilde{X}^2,&\\ [X^2,\tilde{X}^2] + \frac{\partial \tilde{X}^2}{\partial x_1}X^1 - \frac{\partial X^2}{\partial x_1}\tilde{X}^1\bigg)&\end{aligned}$$

In the case of three factors $\mathcal{M}_1\times\mathcal{M}_2\times\mathcal{M}_3$, this generalizes to

$$\begin{aligned}g[X,\tilde{X}] = \bigg(&[X^1,\tilde{X}^1] + \frac{\partial \tilde{X}^1}{\partial x_2}X^2 + \frac{\partial \tilde{X}^1}{\partial x_3}X^3 - \frac{\partial X^1}{\partial x_2}\tilde{X}^2 - \frac{\partial X^1}{\partial x_3}\tilde{X}^3,\\ &[X^2,\tilde{X}^2] + \frac{\partial \tilde{X}^2}{\partial x_1}X^1 + \frac{\partial \tilde{X}^2}{\partial x_3}X^3 - \frac{\partial X^2}{\partial x_1}\tilde{X}^1 - \frac{\partial X^2}{\partial x_3}\tilde{X}^3,\\ &[X^3,\tilde{X}^3] + \frac{\partial \tilde{X}^3}{\partial x_1}X^1 + \frac{\partial \tilde{X}^3}{\partial x_2}X^2 - \frac{\partial X^3}{\partial x_1}\tilde{X}^1 - \frac{\partial X^3}{\partial x_2}\tilde{X}^2\bigg).\end{aligned} \tag{9}$$

If M_i are Euclidean then these specialize to the usual coordinate formula [1].

Suppose $\mathcal{M}$ is a manifold, $\mathcal{D}$ is a distribution of $\mathcal{M}$, and $\mathcal{E}$ is a vector bundle and $\nu\colon \boldsymbol{T}\mathcal{M} \to \mathcal{E}$ is a vector bundle map such that $\mathcal{D} \subseteq \ker \nu$. Then [7] there is a unique $\mathcal{E}$-valued *curvature two form* Δ on $\boldsymbol{T}\mathcal{M}$ such that, for all vector fields $X, \tilde{X} \in \mathcal{D}$,

$$\Delta(X, \tilde{X}) = -\nu[X, \tilde{X}].$$

If $\ker \nu = \mathcal{D}$ then $\mathcal{D}$ is involutive if and only if $\Delta = 0$.

Remember that $\alpha_k = r_k/R$, so $\alpha_1 + \alpha_2 = 1$, the rolling constraint (7) is

$$\dot{u} - (\alpha_1 z_1 \Omega_1 + \alpha_2 z_2 \Omega_2) \times u = 0, \tag{10}$$

and consider the two vector fields (on the configuration space $\mathcal{Q}$)

$$X \equiv (\Omega_1, \Omega_2, U), \qquad \tilde{X} \equiv (\tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{U}),$$

where Ω_k are constant in (z_1, z_2, u), and

$$U = (\alpha_1 z_1 \Omega_1 + \alpha_2 z_2 \Omega_2) \times u, \qquad \tilde{U} = (\alpha_1 z_1 \tilde{\Omega}_1 + \alpha_2 z_2 \tilde{\Omega}_2) \times u,$$

so that the vector fields satisfy the rolling constraint. The required curvature is obtained by substitution of the Lie bracket $[X, \tilde{X}]$ into (10).

Since $\Omega_1, \tilde{\Omega}_1$ and $\Omega_2, \tilde{\Omega}_2$ in X and $\tilde{X}$ represent left invariant vector fields on SO(3), the first and second components of $[X, \tilde{X}]$ are

$$[X, \tilde{X}]^1 = \Omega_1 \times \tilde{\Omega}_1, \qquad [X, \tilde{X}]^2 = \Omega_2 \times \tilde{\Omega}_2.$$

For the third component, as is easily verified, if $b, \tilde{b} \in \mathbb{R}^3$ is constant, then $b \times u$ and $\tilde{b} \times u$ are vector fields on the two sphere with Lie bracket $-(b \times \tilde{b}) \times u$. Also,

$$\frac{\partial U}{\partial z_k} \tilde{\Omega}_k = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} (\alpha_k z_k \exp(\tilde{\Omega}_k^{\wedge} \epsilon) \Omega_k + \alpha_2 z_2 \Omega_2) \times u = \alpha_k z_k (\tilde{\Omega}_k \times \Omega_k) \times u,$$

so, in this way,

$$\begin{aligned} \frac{\partial U}{\partial z_k} \tilde{\Omega}_k - \frac{\partial \tilde{U}}{\partial z_k} \Omega_k &= \alpha_k z_k (\tilde{\Omega}_k \times \Omega_k) \times u - \alpha_k z_k (\Omega_k \times \tilde{\Omega}_k) \times u \\ &= 2\alpha_k z_k (\tilde{\Omega}_k \times \Omega_k) \times u. \end{aligned}$$

Using (9)

$$\begin{aligned} [X, \tilde{X}]^3 = -\big((\alpha_1 z_1 \Omega_1 + \alpha_2 z_2 \Omega_2) \times (\alpha_1 z_1 \tilde{\Omega}_1 + \alpha_2 z_2 \tilde{\Omega}_2)\big) \times u \\ - 2\alpha_1 z_1 (\tilde{\Omega}_1 \times \Omega_1) \times u - 2\alpha_2 z_2 (\tilde{\Omega}_2 \times \Omega_2) \times u, \end{aligned}$$

and substitution into (10) gives

$$\begin{aligned}\Delta &= \big((\alpha_1 z_1\Omega_1 + \alpha_2 z_2\Omega_2) \times (\alpha_1 z_1\tilde{\Omega}_1 + \alpha_2 z_2\tilde{\Omega}_2)\big) \times u \\ &\quad + 2\alpha_1 z_1(\tilde{\Omega}_1 \times \Omega_1) \times u + 2\alpha_2 z_2(\tilde{\Omega}_2 \times \Omega_2) \times u \\ &\qquad + \big(\alpha_1 z_1(\Omega_1 \times \tilde{\Omega}_1) + \alpha_2 z_2(\Omega_2 \times \tilde{\Omega}_2)\big) \times u \\ &= \big(\alpha_1(\alpha_1 - 1) z_1\Omega_1 \times z_1\tilde{\Omega}_1 + \alpha_1\alpha_2\, z_1\Omega_1 \times z_2\tilde{\Omega}_2 \\ &\quad + \alpha_1\alpha_2\, z_2\Omega_2 \times z_1\tilde{\Omega}_1 + \alpha_2(\alpha_2 - 1) z_2\Omega_2 \times z_2\tilde{\Omega}_2\big) \times u \\ &= -\alpha_1\alpha_2 \left(z_1\Omega_1 \times z_1\tilde{\Omega}_1 - z_1\Omega_1 \times z_2\tilde{\Omega}_2 - z_2\Omega_2 \times z_1\tilde{\Omega}_1 + z_2\Omega_2 \times z_2\tilde{\Omega}_2\right) \times u \\ &= \alpha_1\alpha_2\, u \times \big((z_1\Omega_1 - z_2\Omega_2) \times (z_1\tilde{\Omega}_1 - z_2\tilde{\Omega}_2)\big)\end{aligned}$$

Since this curvature is nonzero, the rolling spheres system is nonholonomic.

2.3 Equations of motion

Let $\mathcal{G}$ be a Lie group and let $L(g,\xi)$ be a Lagrangian on the left translated phase space $\boldsymbol{T}\mathcal{G} = \big\{(g,\xi)\big\}$. By Appendix I of [4], the Euler-Lagrange functional is:

$$\delta L = -\frac{d}{dt}\frac{\partial L}{\partial \xi} + \frac{\partial L}{\partial g} \circ \boldsymbol{T_1} L_g + \frac{\partial L}{\partial \xi} \circ \mathrm{ad}_\xi, \tag{11}$$

where L_g denotes left translation. We compute this for the rolling spheres Lagrangian (8) i.e. $G \equiv \mathrm{SO}(3) \times \mathrm{SO}(3) \times \mathbb{R}^3$. The equations of motion result by imposing

1. δL annihilates the rolling constraint; and
2. the rolling constraint itself; and
3. the holonomic constraint $u \cdot u = 1$.

For the case of (8), L does not depend on u, and $\mathbb{R}^3$ is abelian, so the contribution of the u dependence in L is

$$\begin{aligned}&\frac{d}{dt}\frac{\partial L}{\partial \dot{u}} \cdot \big((\alpha_1 z_1\, \delta z_1 + \alpha_2 z_2\, \delta z_2) \times u\big) \\ &\qquad = \alpha_1 z_1^{-1} \left(u \times \frac{d}{dt}\frac{\partial L}{\partial \dot{u}}\right) \cdot \delta z_1 + \alpha_2 z_2^{-1} \left(u \times \frac{d}{dt}\frac{\partial L}{\partial \dot{u}}\right) \cdot \delta z_2.\end{aligned}$$

Also, the third term of (11) gives

$$\frac{\partial L}{\partial \Omega_k} \cdot (\Omega_k \times \delta z_k) = \left(\frac{\partial L}{\partial \Omega_k} \times \Omega_k\right) \cdot \delta z_k$$

and we define $\delta L/\delta z_k \in \mathbb{R}^3$ by

$$\frac{\delta L}{\delta z_1} \cdot \delta z_1 = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L\big(z_1 \exp(\delta z_1\, \epsilon), z_2, u, \Omega_1, \Omega_2, \dot{u}\big),$$

and similarly for $\delta L/\delta z_1$. The result are the equations of motion

$$-\frac{d}{dt}\frac{\partial L}{\partial \Omega_k} - \alpha_k z_k^{-1}\left(u \times \frac{d}{dt}\frac{\partial L}{\partial \dot{u}}\right) + \frac{\delta L}{\delta z_k} + \frac{\partial L}{\partial \Omega_k} \times \Omega_k = 0.$$

The constraint $u \cdot u = 1$ does not affect the variations because the rolling constraint already imposes that $\dot{u} \cdot u = 0$.

In view of (8)

$$\begin{aligned}
\frac{\partial L}{\partial \Omega_1} \cdot \delta\Omega_1 &= \delta\Omega_1^t J_1 \Omega_1 - \dot{u} \cdot \big(\alpha_1 z_1(\delta\Omega_1 \times e_1)\big) \\
&\qquad - \alpha_1\alpha_2\, z_1(\delta\Omega_1 \times e_1) \cdot z_2(\Omega_2 \times e_2) \\
&= \big[J_1\Omega_1 - \alpha_1\alpha_2\, e_1 \times z_1^{-1} z_2(\Omega_2 \times e_2) - \alpha_1 e_1 \times z_1^{-1}\dot{u}\big] \cdot \delta\Omega_1 \\
\frac{\partial L}{\partial \Omega_2} \cdot \delta\Omega_2 &= \delta\Omega_2^t J_2 \Omega_2 + \dot{u} \cdot \big(\alpha_2 z_2(\delta\Omega_2 \times e_2)\big) \\
&\qquad - \alpha_1\alpha_2\, z_1(\Omega_1 \times e_1) \cdot z_2(\delta\Omega_2 \times e_2) \\
&= \big[J_2\Omega_2 - \alpha_1\alpha_2\, e_2 \times z_2^{-1} z_1(\Omega_1 \times e_1) + \alpha_2 e_2 \times z_2^{-1}\dot{u}\big] \cdot \delta\Omega_2 \\
\frac{\partial L}{\partial \dot{u}} \cdot \delta\dot{u} &= \big[\dot{u} - \alpha_1 z_1(\Omega_1 \times e_1) + \alpha_2 z_2(\Omega_2 \times e_2)\big] \cdot \delta\dot{u}
\end{aligned}$$

$$\begin{aligned}
\frac{\delta L}{\delta z_1} \cdot \delta z_1 &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\Big(-\alpha_1 z_1 \exp(\delta z_1^\wedge \epsilon)(\Omega_1 \times e_1) \cdot \dot{u} \\
&\qquad - \alpha_1\alpha_2\, z_1 \exp(\delta z_1^\wedge \epsilon)(\Omega_1 \times e_1) \cdot z_2(\Omega_2 \times e_2)\Big) \\
&= -\alpha_1\big(\delta z_1 \times (\Omega_1 \times e_1)\big) \cdot z_1^{-1}\dot{u} \\
&\qquad - \alpha_1\alpha_2\big(\delta z_1 \times (\Omega_1 \times e_1)\big) \cdot z_1^{-1} z_2(\Omega_2 \times e_2) \\
&= \big[-\alpha_1\alpha_2(\Omega_1 \times e_1) \times z_1^{-1} z_2(\Omega_2 \times e_2) - \alpha_1(\Omega_1 \times e_1) \times z_1^{-1}\dot{u}\big] \cdot \delta z_1 \\
\frac{\delta L}{\delta z_2} \cdot \delta z_2 &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\Big(\alpha_2 z_2 \exp(\delta z_2^\wedge \epsilon)(\Omega_2 \times e_2) \cdot \dot{u} \\
&\qquad - \alpha_1\alpha_2\, z_1(\Omega_1 \times e_1) \cdot z_2 \exp(\delta z_2^\wedge t)(\Omega_2 \times e_2)\Big) \\
&= \alpha_2\big(\delta z_2 \times (\Omega_2 \times e_2)\big) \cdot z_2^{-1}\dot{u} \\
&\qquad - \alpha_1\alpha_2\, z_2^{-1} z_1(\Omega_1 \times e_1) \cdot \big(\delta z_2 \times (\Omega_2 \times e_2)\big)\big] \\
&= \big[-\alpha_1\alpha_2(\Omega_2 \times e_2) \times z_2^{-1} z_1(\Omega_1 \times e_1) + \alpha_2(\Omega_2 \times e_2) \times z_2^{-1}\dot{u}\big] \cdot \delta z_2 \\
\frac{\partial L}{\partial u} &= 0
\end{aligned}$$

The equations of motion become:

$$
\begin{aligned}
&-\frac{d}{dt}\big(J_1\Omega_1 - \alpha_1\alpha_2\, e_1 \times z_1^{-1} z_2(\Omega_2 \times e_2)\big)\\
&+\frac{d}{dt}\big(\alpha_1 e_1 \times z_1^{-1}\dot{u}\big)\\
&-\alpha_1 z_1^{-1}\left(u \times \frac{d}{dt}\big(\dot{u} - \alpha_1 z_1(\Omega_1 \times e_1) + \alpha_2 z_2(\Omega_2 \times e_2)\big)\right)\\
&-\alpha_1\alpha_2(\Omega_1 \times e_1) \times z_1^{-1} z_2(\Omega_2 \times e_2)\\
&-\alpha_1(\Omega_1 \times e_1) \times z_1^{-1}\dot{u}\\
&+\big(J_1\Omega_1 - \alpha_1\alpha_2\, e_1 \times z_1^{-1} z_2(\Omega_2 \times e_2)\big) \times \Omega_1\\
&-\big(\alpha_1 e_1 \times z_1^{-1}\dot{u}\big) \times \Omega_1 = 0.
\end{aligned}
$$

Expand the derivative in the first line, and reorder the lines:

$$
\begin{aligned}
&-\frac{d}{dt}\big(J_1\Omega_1 - \alpha_1\alpha_2\, e_1 \times z_1^{-1} z_2(\Omega_2 \times e_2)\big)\\
&+\alpha_1 e_1 \times z_1^{-1}\frac{d\dot{u}}{dt}\\
&-\alpha_1\big(e_1 \times \big(\Omega_1 \times (z_1^{-1}\dot{u}) + (z_1^{-1}\dot{u}) \times (e_1 \times \Omega_1) + \Omega_1 \times \big((z_1^{-1}\dot{u}) \times e_1\big)\big)\\
&-\alpha_1 z_1^{-1}\left(u \times \frac{d}{dt}\big(\dot{u} - \alpha_1 z_1(\Omega_1 \times e_1) + \alpha_2 z_2(\Omega_2 \times e_2)\big)\right)\\
&-\alpha_1\alpha_2(\Omega_1 \times e_1) \times z_1^{-1} z_2(\Omega_2 \times e_2)\\
&+\big(J_1\Omega_1 - \alpha_1\alpha_2\, e_1 \times z_1^{-1} z_2(\Omega_2 \times e_2)\big) \times \Omega_1 = 0.
\end{aligned}
$$

The third line is zero by the Jacobi identity. Expand the derivative in lines 1 and 4:

$$
\begin{aligned}
&= -J_1\frac{d\Omega_1}{dt}\\
&+\alpha_1(e_1 - z_1^{-1}u) \times z_1^{-1}\frac{d\dot{u}}{dt}\\
&-\alpha_1\alpha_2\, e_1 \times \big(\Omega_1 \times \big(z_1^{-1} z_2(\Omega_2 \times e_2)\big)\big) + \alpha_1\alpha_2\, z_1^{-1} z_2\big(\Omega_2 \times (\Omega_2 \times e_2)\big)\\
&+\alpha_1^2(z_1^{-1}u) \times \big(\Omega_1 \times (\Omega_1 \times e_1)\big) - \alpha_1\alpha_2(z_1^{-1}u) \times z_1^{-1} z_2\big(\Omega_2 \times (\Omega_2 \times e_2)\big)\\
&+\alpha_1^2(z_1^{-1}u) \times \left(\frac{d\Omega_1}{dt} \times e_1\right) - \alpha_1\alpha_2(z_1^{-1}u) \times z_1^{-1} z_2\left(\frac{d\Omega_2}{dt} \times e_2\right)\\
&-\alpha_1\alpha_2(\Omega_1 \times e_1) \times z_1^{-1} z_2(\Omega_2 \times e_2)\\
&+(J_1\Omega_1) \times \Omega_1 - \alpha_1\alpha_2\left(e_1 \times z_1^{-1} z_2(\Omega_2 \times e_2)\right) \times \Omega_1 = 0.
\end{aligned}
$$

Reorder the terms:

$$
\begin{aligned}
&- J_1\frac{d\Omega_1}{dt} + (J_1\Omega_1)\times\Omega_1\\
&+ \alpha_1(e_1 - z_1^{-1}u)\times z_1^{-1}\frac{d\dot u}{dt}\\
&+ \alpha_1^2(z_1^{-1}u)\times\left(\frac{d\Omega_1}{dt}\times e_1\right) - \alpha_1\alpha_2(z_1^{-1}u)\times z_1^{-1}z_2\left(\frac{d\Omega_2}{dt}\times e_2\right)\\
&+ \alpha_1\alpha_2\, z_1^{-1}z_2\big(\Omega_2\times(\Omega_2\times e_2)\big)\\
&+ \alpha_1^2(z_1^{-1}u)\times\big(\Omega_1\times(\Omega_1\times e_1)\big) - \alpha_1\alpha_2(z_1^{-1}u)\times z_1^{-1}z_2\big(\Omega_2\times(\Omega_2\times e_2)\big)\\
&- \alpha_1\alpha_2\, e_1\times\big(\Omega_1\times\big(z_1^{-1}z_2(\Omega_2\times e_2)\big)\big)\\
&- \alpha_1\alpha_2\, z_1^{-1}z_2(\Omega_2\times e_2)\times(e_1\times\Omega_1)\\
&- \alpha_1\alpha_2\,\Omega_1\times\big(z_1^{-1}z_2(\Omega_2\times e_2)\times e_1\big) = 0.
\end{aligned}
$$

The sum of the last three lines is zero by the Jacobi identity. Also, noting that

$$
\begin{aligned}
&\frac{d}{dt}(z_k\Omega_k) = z_k\frac{d\Omega_k}{dt} + z_k(\Omega_k\times\Omega_k) = z_k\frac{d\Omega_k}{dt}\\
&\frac{d\dot u}{dt} = \left(\alpha_1 z_1\frac{d\Omega_1}{dt} + \alpha_2 z_2\frac{d\Omega_2}{dt}\right)\times u + (\alpha_1 z_1\Omega_1 + \alpha_2 z_2\Omega_2)\times\dot u,
\end{aligned}
$$

substitution gives

$$
\begin{aligned}
&- \big(J_1 + \alpha_1^2(e_1 - z_1^{-1}u)^\wedge(z_1^{-1}u)^\wedge + \alpha_1^2(z_1^{-1}u)^\wedge e_1^\wedge\big)\frac{d\Omega_1}{dt}\\
&- \alpha_1\alpha_2\big((e_1 - z_1^{-1}u)^\wedge(z_1^{-1}u)^\wedge - (z_1^{-1}u)^\wedge(z_1^{-1}z_2e_2)^\wedge\big)z_1^{-1}z_2\frac{d\Omega_2}{dt}\\
&+ (J_1\Omega_1)\times\Omega_1 + \alpha_1^2(z_1^{-1}u)\times\big(\Omega_1\times(\Omega_1\times e_1)\big)\\
&+ \alpha_1\alpha_2\, z_1^{-1}z_2\big(\mathbf{1} - (z_2^{-1}u)^\wedge\big)\Omega_2\times(\Omega_2\times e_2)\\
&+ \alpha_1(e_1 - z_1^{-1}u)\times\big((\alpha_1\Omega_1 + \alpha_2 z_1^{-1}z_2\Omega_2)\times(z_1^{-1}\dot u)\big) = 0
\end{aligned}
$$

A second equation can be obtained from this by exchange of the indices $k = 1$ and $k = 2$, and negation of u.

2.4 Holonomic subsystem

Holonomic and nonholonomic systems are structurally different. For example, nonholonomic systems are not symplectic, there is not a strict Noether correspondence between symmetry and momentum, and there can be attractive behavior near equilibria. However, mere failure of a system to be holonomic does not necessarily imply the absence of any holonomic property:

Theorem ([7], Theorem 5.2)**.** *Let $(\mathcal{P}, \omega, \mathcal{K}, H)$ be a semi-Hamiltonian system, and let $\mathcal{K}_0$ be a subbundle of $\mathcal{K}$ over $\mathcal{P}_0 \subseteq \mathcal{P}$. Suppose that*

1. *$\mathcal{K}_0$ is $\boldsymbol{T}F_t$ invariant, where F_t is the flow of Y_H;*

2. $\Delta_{\mathcal{K}} = 0$ *on* $\mathcal{K}_0$*;*

3. $\boldsymbol{d}\omega = 0$ *on* $\mathcal{K}_0$*.*

Then F_t *is symplectic on* $\mathcal{K}_0$*.*

The theorem is stated in the semisymplectic category, which supposes a distribution $\mathcal{K}$ on phase space and a nondegenerate two form ω on $\mathcal{K}$. The semisymplectic evolution is provided by the semi-Hamiltonian vector field, which is defined by $\boldsymbol{i}_Y\omega = \boldsymbol{d}H$ on $\mathcal{K}$ with Y taking values in $\mathcal{K}$. The relation to the Lagrangian formulation is (1) $\mathcal{K} = \hat{\mathcal{D}} \equiv (\boldsymbol{T}\tau_{\mathcal{Q}})^{-1}\mathcal{D} \cap \boldsymbol{T}\mathcal{D}$, where $\tau_{\mathcal{Q}} \colon \boldsymbol{T}\mathcal{Q} \to \mathcal{Q}$ is the canonical projection, and (2) the semisymplectic form is the Lagrange two form ω_L; see [7] for more details and references. We will show that the rolling disk system occurs as a holonomic subsystem of the rolling spheres system.

Let $\mathrm{SO}(3)_3 \equiv \left\{ A \in \mathrm{SO}(3) : A\mathbf{e}_3 = \mathbf{e}_3 \right\}$, where $\mathbf{e}_3 = (0,0,1)$, define

$$\mathcal{Q}_0 \equiv \mathrm{SO}(3)_3 \times \mathrm{SO}(3)_3 \times S^2,$$

assume $e_1 \cdot \mathbf{e}_3 = 0, e_2 \cdot \mathbf{e}_3 = 0$, and assume that J_k are diagonal. Such configurations and parameters correspond to the two spheres with centers of mass and contact point all in the plane through the origin and perpendicular to $\mathbf{e}_3$. One expects $\boldsymbol{T}\mathcal{Q}_0$ to be an invariant submanifold on which the system is the same as the rolling disk system. In any case the restricted system is holonomic because the curvature Δ is zero on $\boldsymbol{T}\mathcal{Q}_0$, as there it involves the cross product of vectors parallel to $\mathbf{e}_3$.

The invariance of $\boldsymbol{T}\mathcal{Q}_0$ under the flow of the rolling spheres system, could be verified by showing that the evolution vector field of the rolling spheres system is tangent to $\boldsymbol{T}\mathcal{Q}_0$. But the evolution vector field is algebraically complicated, so we verify, directly using the relevant variational principles, that the critical curves on $\mathcal{Q}_0$ are critical in the full space $\mathcal{Q}$.

Let $z_1(t), z_2(t), u(t) \in \mathcal{Q}_0$ be a critical curve of the rolling spheres variational principle, with the additional (holonomic) constraint $\mathcal{Q}_0$. Consider variations of the form $z_k(t)\exp\bigl(\epsilon\delta z_k(t)^\wedge\bigr)$, where $\delta z_k(t)\cdot\mathbf{e}_3 = 0$, and variations $u(t,\epsilon)$, where

$$\delta u \equiv \frac{d}{d\epsilon}\bigg|_{\epsilon=0} u(t,\epsilon)$$

satisfies

$$\delta u = (\alpha_1 z_1\,\delta z_1 + \alpha_2 z_2\,\delta z_2) \times u.$$

It suffices to show that $\boldsymbol{d}L$ annihilates such variations pointwise, because the curve $z_1(t), z_2(t), u(t) \in \mathcal{Q}_0$ is assumed critical for the disk system, and every variation is the sum of variations tangent to $\mathcal{Q}_0$ and with variations δz_k perpendicular to $\mathbf{e}_3$.

Actually, each of the five terms in (8) separately annihilate such variations pointwise: For the terms $\frac{1}{2}\Omega_k^t J_k \Omega_k$, compute Ω_k and $\delta\Omega_k$ as follows

$$\begin{aligned}\Omega_k^\wedge &= \big(z_k(t)\exp\big(\epsilon\,\delta z_k(t)^\wedge\big)\big)^{-1}\frac{d}{dt}\big(z_k(t)\exp\big(\epsilon\,\delta z_k(t)^\wedge\big)\big)\\ &= \exp\big(-\epsilon\,\delta z_k(t)^\wedge \Omega_k\big)^\wedge + \epsilon\,\delta z_k(t)^\wedge,\end{aligned}$$

and so $\delta\Omega_k = -\delta z_k \times \Omega_k + \delta z_k$. Since $\Omega_k \parallel \mathbf{e}_3$ and $\delta z_k \perp \mathbf{e}_3$, we conclude that $\delta\Omega_k \perp \mathbf{e}_3$ and

$$\delta\left(\frac{1}{2}\Omega_k^t J_k \Omega_k\right) \equiv \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\frac{1}{2}\Omega_k^t J_k \Omega_k = \delta\Omega_k \cdot J_k\Omega_k = 0.$$

Similarly, for the last term of (8),

$$\delta\big(z_k(\Omega_k \times e_k)\big) = z_k\big(\delta z_k \times (\Omega_k \times e_k)\big) + z_k(\delta\Omega_k \times e_k).$$

Both terms are parallel to $\mathbf{e}_3$ because $\delta z_k, \delta\Omega_k, e_k \perp \mathbf{e}_3$ and $\Omega_k \parallel \mathbf{e}_3$. For the same reasons, $z_k(\Omega_k \times e_k) \perp \mathbf{e}_3$, so

$$\begin{aligned}&\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0} z_1(\Omega_1 \times e_1)\cdot z_2(\Omega_2 \times e_2) =\\ &\qquad \delta\big(z_1(\Omega_1 \times e_1)\big)\cdot z_2(\Omega_2 \times e_2) + z_1(\Omega_1 \times e_1)\cdot \delta\big(z_1(\Omega_1 \times e_1)\big) = 0.\end{aligned}$$

For the term $\frac{1}{2}|\dot{u}|^2$, $\dot{u} = \frac{\partial}{\partial t}u(t,\epsilon)$ and at $\epsilon = 0$ this is perpendicular to $\mathbf{e}_3$, whereas

$$\begin{aligned}\delta\dot{u} &= \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\frac{\partial}{\partial t}u(t,\epsilon) = \frac{\partial}{\partial t}\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}u(t,\epsilon) = \frac{\partial}{\partial t}\big((\alpha_1 z_1\,\delta z_1 + \alpha_2 z_2\,\delta z_2)\times u\big)\\ &= \frac{\partial}{\partial t}\big(\alpha_1 z_1\,\delta z_1 + \alpha_2 z_2\,\delta z_2\big)\times u\\ &\qquad + (\alpha_1 z_1\,\delta z_1 + \alpha_2 z_2\,\delta z_2)\times\dot{u}.\end{aligned}$$

In this expression the first factor of the first term is the derivative of a curve perpendicular to $\mathbf{e}_3$ and the second factor of the first term is also perpendicular to $\mathbf{e}_3$, so the cross product is parallel to $\mathbf{e}_3$. Similarly, the second term is parallel to $\mathbf{e}_3$. Thus $\delta(|\dot{u}|^2) = \dot{u}\cdot\delta\dot{u} = 0$. For the fourth term in (8)

$$\begin{aligned}&\delta\dot{u}\cdot\big(\alpha_1 z_1(\Omega_1 \times e_1) - \alpha_2 z_2(\Omega_2 \times e_2)\big)\\ &\qquad + \dot{u}\cdot\delta\big(\alpha_1 z_1(\Omega_1 \times e_1) - \alpha_2 z_2(\Omega_2 \times e_2)\big),\end{aligned}$$

and $\delta\dot{u} \parallel \mathbf{e}_3$ whereas $z_k(\Omega_k \times e_k) \perp \mathbf{e}_3$. For the second term, $\dot{u} \perp \mathbf{e}_3$ and $\delta\big(\alpha_k z_k(\Omega_k \times e_k)\big) \parallel \mathbf{e}_3$, as derived above.

References

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison-Wesley, second edition, 1978.

[2] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, and R. M. Murray. Nonholonomic systems with symmetry. *Arch. Rational Mech. Anal.*, 136:21–99, 1996.

[3] A. M. Bloch, J. E. Marsden, and D. V. Zenkov. Nonholonomic dynamics. *Notices Amer. Math. Soc.*, 52:324–333, 2005.

[4] J. Hermans. *Rolling rigid bodies with and without symmetries*. PhD thesis, University of Utrecht, 1995.

[5] G. W. Patrick. *Two axially symmetric coupled rigid bodies: relative equilibria, stability, bifurcations, and a momentum preserving symplectic integrator*. PhD thesis, University of California at Berkeley, 1991.

[6] G. W. Patrick. Variational proof that relative equilibria are critical points of the amended Lagrangian. *Differ. Geom. Appl.*, 24:479–481, 2006.

[7] G. W. Patrick. Variational development of the geometry of nonholonomic mechanics. *Rep. Math. Phys.*, 59:145–184, 2007.