Fields Institute Communications
Volume 7, 1996

Optimal Control of a Rigid Body with Two Oscillators

R. Yang, P. S. Krishnaprasad, and W. Dayawansa
Department of Electrical Engineering
and
Institute for Systems Research
University of Maryland
College Park, MD 20742, USA

Abstract This paper is concerned with the exploration of reduction and explicit solvability of optimal control problems on principal bundles with connections from a Hamiltonian point of view. The particular mechanical system we consider is a rigid body with two driven oscillators, for which the bundle structure is $(SO(3) \times \mathbb{R}^2, \mathbb{R}^2, SO(3))$. The optimal control problem is posed by considering a special nonholonomic variational problem, in which the nonholonomic distribution is defined via a connection. The necessary conditions for the optimal control problem are determined intrinsically by a Hamiltonian formulation. The necessary conditions admit the structure group of the principal bundle as a symmetry group of the system. Thus the problem is amenable to Poisson reduction. Under suitable hypotheses and approximations, we find that the reduced system possesses additional symmetry which is isomorphic to S^1. Applying Poisson reduction again, we obtain a further reduced system and corresponding first integral. These reductions imply explicit solvability for suitable values of parameters.

1 Introduction and Background

An interesting problem in multibody mechanics is the problem of nonholonomic motion planning, or the kinematic control problem. In recent research on various multibody mechanical systems with symmetry, the theory of principal bundles with connections has lead to clear insight into the geometric structure of the problem, and provided a common framework for the formulation of related optimal control problems. However, explicit or partially explicit solution to the necessary conditions, given by differential equations on phase space, for the optimal path and

---

1991 *Mathematics Subject Classification.* Primary 49J15; Secondary 55R10, 70E15, 70H99.

This work was supported in part by the AFOSR University Research Initiative Program under grant AFOSR-90-0105, by the National Science Foundation's Engineering Research Centers Program: NSFD CDR 8803012, and by the National Science Foundation Grant ECS-909612.

control is still a challenge. Although, under certain conditions, the symmetries of the systems imply the existence of conserved quantities for the differential equations given by the necessary conditions, working with local coordinates at the early stage of the analysis usually causes difficulties in discovering such quantities. In this paper, we consider a particular mechanical system consisting of a rigid body and two point-masses, for which the structure group of the principal bundle is non-Abelian in general. We will formulate the related optimal control problem and the corresponding necessary conditions intrinsically from a Hamiltonian perspective and explore their explicit solvability. Some results in this paper have been presented in our previous work [1]. Here, we provide detailed proofs of those results.

As in [1], the kinematic control problem considered is based on the following abstract geometric objects. Consider a *simple mechanical system with symmetry* (following terminology of Smale [2]), (Q, K, V, G), where the configuration space Q is a Riemannian manifold with metric K and the Lie group G acts freely on Q on the *left* by isometries and leaves the potential energy V invariant. The action of G on Q is denoted by Φ. Together with this system is an equivariant momentum map with respect to the tangential action Φ^T of G on TQ, $\mathbf{J} : TQ \to \mathcal{G}^*$ satisfying

$$\langle \mathbf{J}(q,v),\ \xi\rangle = (K^\flat v_q)(\xi_Q(q)) = K(q)(v_q, \xi_Q(q)), \qquad \forall \xi \in \mathcal{G}, \tag{1.1}$$

where $\mathcal{G}^*$ is the dual of the Lie algebra, denoted by $\mathcal{G}$, of G and $K^\flat$ is the standard Legendre transform. In addition, we also consider Q as the total space of a principal fiber bundle, $\wp = (Q, B, \pi, G)$, where $B = Q/G$ is called the base (or shape) space and $\pi : Q \to B$ is the bundle projection. On this bundle, the mechanical connection is constructed as follows. At each point $q \in Q$, define the *locked inertia tensor* as the mapping

$$\mathbb{I}(q) : \mathcal{G} \to \mathcal{G}^* \tag{1.2a}$$

such that

$$\langle \mathbb{I}(q)\eta,\ \xi\rangle = K(q)(\eta_Q(q), \xi_Q(q)) \qquad \forall \eta, \xi \in \mathcal{G}. \tag{1.2b}$$

Then, the *mechanical connection* is defined by the $\mathcal{G}$-valued one-form:

$$\alpha : TQ \to \mathcal{G}\colon (q,v) \mapsto \alpha(q,v) = \mathbb{I}^{-1}(q)\mathbf{J}(q,v). \tag{1.3}$$

Indeed, one can show that $\alpha(\xi_Q(q)) = \xi, \forall \xi \in \mathcal{G}$ and

$$(\Phi_g^* \alpha)(q,v) = \mathrm{Ad}_g\, \alpha(q,v), \qquad \forall g \in G.$$

The mechanical connection appears to be originally due to Smale and Kummer (see [3]). With this well-defined connection, we have a vertical-horizontal splitting of the tangent bundle TQ,

$$T_qQ = (\mathrm{Ver})_q \oplus (\mathrm{Hor})_q \tag{1.4a}$$

such that, for each $v_q \in T_qQ$,

$$\begin{aligned} v_q &= (\alpha(v_q))_Q(q) + (v_q - (\alpha(v_q))_Q(q)) \\ &= (\mathbb{I}(q)^{-1}\mu)_Q(q) + (v_q - (\mathbb{I}(q)^{-1}\mu)_Q(q)), \end{aligned} \tag{1.4b}$$

where $\mu = \mathbf{J}(q, v)$. It is readily shown that

$$\text{Hor} = \{(q, v) \in TQ \mid \mathbf{J}(q, v) = 0\}, \tag{1.5}$$

and the splitting in (1.4) is the orthogonal one with respect to metric K.

To formulate the kinematic control problem explicitly, we consider the trivial bundle, i.e., $\wp = (B \times G, B, \pi, G)$. Here, the control is *internal* to the system, which leaves invariant the conserved momentum map $\mathbf{J}$. Since, by definition, a principal fiber bundle is locally trivial, the equations we have below are locally true in general.

Represent the tangent space at each point $(x, g) \in Q$ by

$$T_{(x,g)}Q = T_x B \times T_g G$$

and let a tangent vector in $T_{(x,g)}Q$ be represented by

$$v_{(x,g)} = (\dot{x}, \dot{g})_{(x,g)} = (\dot{x}, g\xi)_{(x,g)},$$

where $\xi(t) = T_g L_{g^{-1}} \dot{g} \in \mathcal{G}$. The Lie group G acts on Q following the rule $\Phi(h, (x, g)) = (x, hg)$, where $h, g \in G$ and $x \in B$. Then the infinitesimal generator corresponding to $\eta \in \mathcal{G}$ is

$$\eta_Q(q) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \Phi(\exp(\epsilon\eta), (x, g)) = (0, \eta \cdot g). \tag{1.6}$$

Using the G-invariance of K, we have

$$\begin{aligned}
\mathbf{J}(q, v) \cdot \eta &= K(x, g)((\dot{x}, g\xi), (0, \eta \cdot g)) \\
&= K(x, e)((\dot{x}, \xi), (0, \mathrm{Ad}_{g^{-1}} \eta)) \\
&= \langle \widetilde{\mathbb{I}}(x)\xi,\ \mathrm{Ad}_{g^{-1}} \eta\rangle + \langle \jmath(x)(\dot{x}),\ \mathrm{Ad}_{g^{-1}} \eta\rangle \\
&= \langle \mathrm{Ad}^*_{g^{-1}} (\widetilde{\mathbb{I}}(x)\xi + \jmath(x)(\dot{x})),\ \eta\rangle,
\end{aligned}$$

where e is the identity element in G; $\widetilde{\mathbb{I}}(x) \triangleq \mathbb{I}(x, e)$ is referred to as the (local) locked inertia tensor at x, and represents the metric on G, and $\jmath(x) : T_x B \to \mathcal{G}^*$ comes from the cross term when the metric K is written in terms of metrics on B and G. Here, the metric on B is induced from K. Therefore, we have, for $\mu = \mathbf{J}(q, v)$,

$$\mu = \mathrm{Ad}^*_{g^{-1}} \widetilde{\mathbb{I}}(x)\xi + \mathrm{Ad}^*_{g^{-1}} \jmath(x)(\dot{x}) \tag{1.7}$$

or

$$\xi = \widetilde{\mathbb{I}}(x)^{-1} \mathrm{Ad}^*_g \mu - \widetilde{\mathbb{I}}(x)^{-1} \jmath(x)\dot{x}$$

or, by left action,

$$\dot{g}(t) = g(t) \cdot (\widetilde{\mathbb{I}}(x)^{-1} \mathrm{Ad}^*_g \mu - \widetilde{\mathbb{I}}(x)^{-1} \jmath(x)\dot{x}). \tag{1.8}$$

Given a closed curve in B and an initial point $q_0 = (x_0, g_0)$ in Q, using (1.8) one will be able to compute the shift, or the *phase*, in G. The phase generated by the first term in (1.8) is referred to as a *dynamic phase* and the phase generated by the second term in (1.8) is the holonomy, referred to as *geometric phase*. One can show that $\widetilde{\mathbb{I}}(x)^{-1}\jmath(x)(\dot{x})$ is, in fact, the value of the local connection form of the mechanical connection at $(x, \dot{x})$.

Assuming that the vector $\dot{x}$ or the velocity of the path in B can be directly controlled, from (1.8), an associated kinematic control system can be set up as

$$\begin{cases} \dot{x} = u, \\ \dot{g} = g \cdot \left(\widetilde{\mathbb{I}}(x)^{-1} \operatorname{Ad}_g^* \mu - \widetilde{\mathbb{I}}(x)^{-1} \jmath(x) u\right), \end{cases} \tag{1.9a}$$

or simply

$$\dot{q} = X_\mu(q) + \mathcal{H}(q)u, \tag{1.9b}$$

for $q = (g, x) \in Q$, where $X_\mu(q) = (0, g \cdot \widetilde{\mathbb{I}}(x)^{-1} \operatorname{Ad}_g^* \mu)$ is the drift, $\mathcal{H}(q) : T_{\pi(q)}(B) \to T_q Q$ is the horizontal lift operator and $u \in T_{\pi(q)}(B)$ is a tangent vector on shape space representing controls. Two control problems for this system can be framed as follows:

(P1) Given two points q_0 and q_1 in Q, find $u(\cdot)$ steering q_0 to q_1 at a specified time;

(P2) Given two points q_0 and q_1 in Q on the same fiber, find $u(\cdot)$ steering q_0 to q_1 while minimizing

$$\int_0^T \langle u, u \rangle_B \, dt$$

for the Riemannian metric $\langle \cdot, \cdot \rangle_B$ on B and the fixed final time $T > 0$ subject to (1.9).

The problems (P1) and (P2) are standard problems in control theory, namely controllability and optimal control. If $\mu = 0$, the problem of controllability is settled by appealing to Chow's theorem [4]. In addition, if $\mu = 0$ and the system is controllable, (P2) is the isoholonomic problem in [5], or a special case of the problem of singular Riemannian/sub-Riemannian/nonholonomic geodesics [6, 7].

In the next section, we will formulate the control system and corresponding optimal control problem for the system of a rigid body with two oscillators following the above procedure.

2 Mechanical Connection for the System

In this section, we give an explicit expression for the mechanical connection for the system consisting of a rigid body with two (driven) oscillators.

The mechanical system we consider is shown in Figure 2.1. Here, r_0 is the position vector of the center of mass of the rigid body or *carrier* relative to the center of mass of the system; r_1 and r_2 are the position vectors of two oscillators with point masses m_1 and m_2 relative to the center of mass of the system, respectively; the mass and moment of inertia tensor of the carrier are denoted as m_0 and I_0, respectively; Q_1 and Q_2 are the position vectors of two oscillators relative to a frame (not displayed) fixed on the carrier. We assume that no exterior force/torque affects

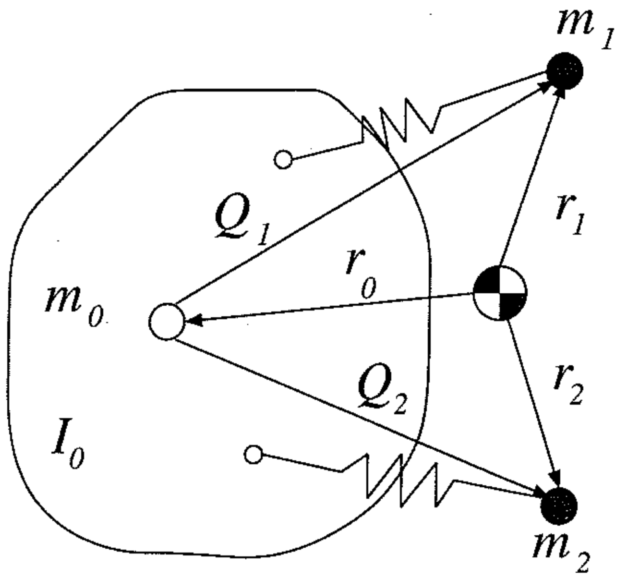

Figure 2.1 A Rigid Body with Two Oscillators

the system and the potential energy V is zero. The inertial frame can be placed at the center of mass of the system and r_0, r_1 and r_2 are related by

$$\sum_{i=0}^{2} m_i r_i = 0. \tag{2.1}$$

For now, r_1 and r_2 (or Q_1 and Q_2) are assumed to be arbitrary time dependent vectors. Later, we will impose constraints on them to study the effect of their motion on the motion of the carrier.

From the above setting, we have the configuration space $Q = (\mathbb{R}^3)^2 \times SO(3)$ with coordinates $q = (r_1, r_2, A)$ and its tangent bundle $TQ = (T\mathbb{R}^3)^2 \times TSO(3)$ with local coordinates $(q, v) = ((r_1, r_2, A), (\dot{r}_1, \dot{r}_2, A\widehat{\Omega}))$. Here, denoting by Ω the vector of angular velocity of the carrier with respect to the body fixed frame, we used the fact

$$\dot{A} = A\widehat{\Omega} \tag{2.2}$$

with the standard isomorphism

$$\widehat{\cdot} : \mathbb{R}^3 \to so(3) : \ (x_1, x_2, x_3) \mapsto \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}.$$

The Lagrangian of the system can be determined easily as

$$\begin{aligned} L((r_1,r_2,A),(\dot r_1,\dot r_2,A\widehat\Omega)) =& \tfrac{1}{2}\langle\Omega, I_0\Omega\rangle + \tfrac{1}{2}\Big(m_1+\frac{m_1^2}{m_0}\Big)\langle\dot r_1,\dot r_1\rangle \\ &+\tfrac{1}{2}\Big(m_2+\frac{m_2^2}{m_0}\Big)\langle\dot r_2,\dot r_2\rangle + \frac{m_1m_2}{m_0}\langle\dot r_1,\dot r_2\rangle, \end{aligned} \tag{2.3}$$

where $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^3$. This Lagrangian is given by a Riemannian metric K on Q, i.e.,

$$L(q,v) = \tfrac{1}{2}K(q)(v,v),$$

where, for $(u_1,u_2,A\widehat\Omega)$ and $(w_1,w_2,A\widehat\Xi)\in T_{(r_1,r_2,A)}Q$,

$$\begin{aligned} &K(r_1,r_2,A)((u_1,u_2,A\widehat\Omega),(w_1,w_2,A\widehat\Xi)) \\ &\triangleq \langle\Omega,I_0\Xi\rangle + \Big(m_1+\frac{m_1^2}{m_0}\Big)\langle u_1,w_1\rangle + \Big(m_2+\frac{m_2^2}{m_0}\Big)\langle u_2,w_2\rangle \\ &\quad + \frac{m_1m_2}{m_0}\langle u_1,w_2\rangle + \frac{m_1m_2}{m_0}\langle u_2,w_1\rangle. \end{aligned} \tag{2.4}$$

Let $G=SO(3)$ act on Q by

$$\begin{aligned} \Phi: SO(3)\times((\mathbb{R}^3)^2\times SO(3)) &\to (\mathbb{R}^3)^2\times SO(3), \\ (A,(r_2,r_2,B)) &\mapsto (Ar_1,Ar_2,AB). \end{aligned} \tag{2.5}$$

From (2.4), one can show that G acts on Q by isometries. Therefore, the system $(Q=(\mathbb{R}^3)^2\times SO(3),K,V=0,G=SO(3))$ is a simple mechanical system with symmetry.

By standard intrinsic calculations on $SO(3)$ (cf. [8]), one finds the Legendre transform, at $(q,v)\in TQ$, as

$$\begin{aligned} K^\flat(q)(v) &= D_2L(q,v) \\ &= \Big(A\widehat{I_0\Omega}, \Big(m_1+\frac{m_1^2}{m_0}\Big)\dot r_1 + \frac{m_1m_2}{m_0}\dot r_2, \Big(m_2+\frac{m_2^2}{m_0}\Big)\dot r_2 + \frac{m_1m_2}{m_0}\dot r_1\Big). \end{aligned} \tag{2.6}$$

The infinitesimal generator of the action in (2.5) corresponding to $\widehat\xi\in so(3)$ is

$$\xi_Q(q) = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\Phi(e^{\epsilon\widehat\xi},(r_1,r_2,A)) = (\widehat\xi r_1,\widehat\xi r_2,\widehat\xi A). \tag{2.7}$$

Then, using (1.1), (2.6) and (2.7), the momentum map can be written explicitly as, for $\mu = \mathbf{J}((r_1,r_2,A),(\dot r_1,\dot r_2,A\widehat\Omega))$,

$$\begin{aligned} \mu = AI_0\Omega &+ \Big(m_1+\frac{m_1^2}{m_0}\Big)r_1\times\dot r_1 + \frac{m_1m_2}{m_0}r_1\times\dot r_2 \\ &+ \Big(m_2+\frac{m_2^2}{m_0}\Big)r_2\times\dot r_2 + \frac{m_1m_2}{m_0}r_2\times\dot r_1. \end{aligned} \tag{2.8}$$

One can show that μ is, in fact, the total angular momentum of the system.

It is clear that $((\mathbb{R}^3)^2\times SO(3))$ is a trivial bundle with the structure group $SO(3)$ and the base space $(\mathbb{R}^3)^2$ coordinatized by (Q_1,Q_2). Using coordinates (Q_1,Q_2,A)

for the configuration space Q, the angular momentum (2.8) of the system can be rewritten as follows. From Figure 2.1 and Equation (2.1), we have

$$r_i = r + AQ_i, \qquad i = 1, 2$$

and

$$r = -A(\epsilon_1 Q_1 + \epsilon_2 Q_2),$$

where

$$\epsilon_i = \frac{m_i}{m_0 + m_1 + m_2}, \qquad i = 1, 2.$$

Equation (2.8) can be rearranged as

$$\mu = A((I_0 + \Delta I_0)\Omega + D_1\dot{Q}_1 + D_2\dot{Q}_2), \tag{2.9}$$

where

$$\begin{aligned} \Delta I_0 &= -m(\epsilon_1 \widehat{Q}_1^2 + \epsilon_2 \widehat{Q}_2^2 - (\epsilon_1 \widehat{Q}_1 + \epsilon_2 \widehat{Q}_2)^2) \\ D_1 &= m[(1-\epsilon_1)\epsilon_1 \widehat{Q}_1 - \epsilon_1\epsilon_2 \widehat{Q}_2] \\ D_2 &= m[-\epsilon_1\epsilon_2 \widehat{Q}_1 + \epsilon_2(1-\epsilon_2)\widehat{Q}_2]. \end{aligned}$$

By (2.9), we have

$$\Omega = (I_0 + \Delta I_0)^{-1}(A^T\mu - (D_1\dot{Q}_1 + D_2\dot{Q}_2)) \tag{2.10}$$

or, by (2.2),

$$\dot{A} = A[\,(I_0 + \Delta I_0)^{-1}A^T\mu - (I_0 + \Delta I_0)^{-1}(D_1\dot{Q}_1 + D_2\dot{Q}_2)]^{\frown}, \tag{2.11}$$

where $[\cdot]^{\frown} \triangleq \widehat{(\cdot)}$. Comparing (2.11) with (1.8), we see that

$$I_{\text{lock}}(Q_1, Q_2) \triangleq I_0 + \Delta I_0$$

is the (local) locked inertia tensor, and

$$\omega(Q_1, Q_2)(\dot{Q}_1, \dot{Q}_2) \triangleq [(I_0 + \Delta I_0)^{-1}(D_1\dot{Q}_1 + D_2\dot{Q}_2)]^{\frown}$$

is the value of the (local) connection form at the point $((Q_1, Q_2), (\dot{Q}_1, \dot{Q}_2)) \in TB$ with respect to the mechanical connection. This connection form can be explicitly given by

$$\begin{aligned} \check{\omega}(Q_1, Q_2) &= I_{\text{lock}}^{-1}(D_1 dQ_1 + D_2 dQ_2) \\ &= mI_{\text{lock}}^{-1}[((1-\epsilon_1)\epsilon_1 \widehat{Q}_1 - \epsilon_1\epsilon_2 \widehat{Q}_2)dQ_1 \\ &\qquad + (-\epsilon_1\epsilon_2 \widehat{Q}_1 + \epsilon_2(1-\epsilon_2)\widehat{Q}_2)dQ_2], \end{aligned} \tag{2.12}$$

where the operator $\check{}$ is the inverse of operator $\widehat{}$.

Equation (2.11) can be used for computing the phase of the system and the related optimal control problem mentioned in the preceding section. In particular, when $\mu = 0$, it can be used to compute holonomy, or geometric phase and to solve the isoholonomy problem. In this case, the angular velocity vector of the rigid body in a body fixed frame relates to the connection form by

$$\begin{aligned}\Omega &= -\check{\omega}(Q_1, Q_2)(\dot{Q}_1, \dot{Q}_2) \\ &= mI_{\text{lock}}^{-1}[((1-\epsilon_1)\epsilon_1\widehat{Q}_1 - \epsilon_1\epsilon_2\widehat{Q}_2)\dot{Q}_1 \\ &\qquad + (-\epsilon_1\epsilon_2\widehat{Q}_1 + \epsilon_2(1-\epsilon_2)\widehat{Q}_2)\dot{Q}_2].\end{aligned} \tag{2.13}$$

In the following sections, we will consider the case $\mu = 0$ only. In addition, we will assume that the oscillators are confined to move along certain guide ways. Under this assumption, the bundle structure will be simplified and the equations for phases and the connection form on such a bundle can be easily derived from those we have found.

3 Planar System

We now assume that the vectors r, BQ_1 and BQ_2 are kept in the same plane in inertial space and that $m_1 = m_2$ (so $\epsilon_1 = \epsilon_2 \triangleq \epsilon$). In addition, we choose a body-fixed coordinate system (or frame) 0-xyz on the carrier with 0-z axis perpendicular to the plane and the origin of this frame is placed at the center of mass of the carrier (see Figure 3.1). We also assume that the two oscillators move along two parallel guide ways such that, in the 0-xyz frame, at each time t,

$$Q_1 = (-l, x_1(t), 0)^T \quad \text{and} \quad Q_2 = (\ l, x_2(t), 0)^T,$$

where $x_1(t), x_2(t) \in \mathbb{R}$ and l is the distance of the guide ways to the origin of 0-xyz frame. It is clear that the configuration space is reduced to $Q = \mathbb{R}^2 \times S^1$, which will be coordinatized by (x_1, x_2, θ), and the principal fiber bundle is $(\mathbb{R}^2 \times S^1, \mathbb{R}^2, \pi, S^1)$.

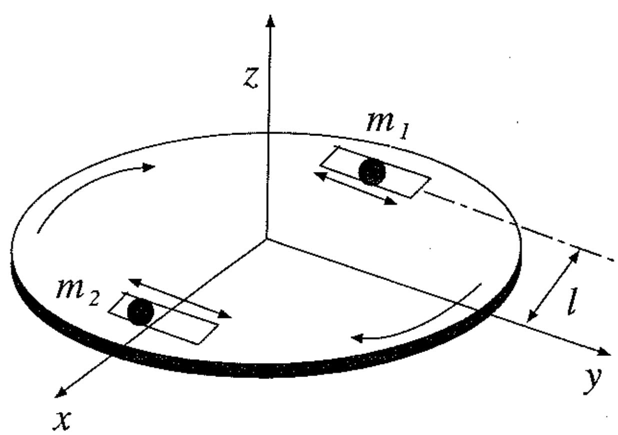

Figure 3.1 A Rigid Body with Two Oscillators: Planar Case

Setting $\mu = 0$, the angular velocity $\Omega = (\Omega_x, \Omega_y, \Omega_z)^T$ in (2.13) is of the form

$$\begin{cases} \Omega_x = 0, \\ \Omega_y = 0, \\ \Omega_z = \dot{\theta} = \dfrac{m\epsilon l}{I_{\text{lock}}}(\dot{x}_1 - \dot{x}_2), \end{cases} \tag{3.1}$$

where

$$I_{\text{lock}} \triangleq I_z + m\epsilon(2l^2 + (1-\epsilon)x_1^2 - 2\epsilon x_1 x_2 + (1-\epsilon)x_2^2),$$

with I_z the moment of inertia of the carrier about z axis. It is obvious that the local connection form corresponding to the mechanical connection on the principal bundle $(\mathbb{R}^2 \times S^1, \mathbb{R}^2, S^1)$ is

$$\omega(x_1, x_2) = -\frac{m\epsilon l}{I_{\text{lock}}}(dx_1 - dx_2). \tag{3.2}$$

For simplicity, we further assume that the amplitude of motion of each oscillator is very small in comparison with the the spacing of two guide ways, i.e.,

$$|x_i|/l \ll 1. \tag{3.3}$$

Under this assumption, using Taylor expansion (up to quadratic terms of x_i), we get an approximate ω (with the same notation)

$$\omega = -\frac{m\epsilon l(I_z + 2m\epsilon l^2 - m\epsilon(2l^2 + (1-\epsilon)x_1^2 - 2\epsilon x_1 x_2 + (1-\epsilon)x_2^2))}{(I_z + 2m\epsilon l^2)^2}(dx_1 - dx_2).$$

The above procedure is called *localization* in [9]. Since we are interested in the motion of the carrier generated by the motion of the point masses on a *closed curve* in shape space, the above ω can be further simplified as follows. Applying the exterior derivative to the above equation, we have

$$d\omega = -\frac{m\epsilon l}{(I_z + 2m\epsilon l^2)^2} d(x_1^2 dx_2 - x_2^2 dx_1).$$

Then, under the assumption (3.3), we can take (for closed paths in shape space)

$$\omega = -\frac{m\epsilon l}{(I_z + 2m\epsilon l^2)^2}(x_1^2 dx_2 - x_2^2 dx_1) \tag{3.4}$$

modulo an exact one-form.

Let $c(\cdot)$ be any closed curve in shape space $\mathbb{R}^2$. Since S^1 is Abelian, from (1.8), the corresponding geometric phase or holonomy, i.e., the drift of the carrier about the z-axis, will be

$$\delta_z = -\int_c \omega = -\int_c \frac{m\epsilon l}{(I_z + 2m\epsilon l^2)^2}(x_1^2 dx_2 - x_2^2 dx_1). \tag{3.5}$$

Using (3.5), we now compute the geometric phase for the case in which both oscillators follow sinusoidal motions with different amplitudes, frequencies and phase angles.

Let

$$x_1(t) = a_1 \sin(\bar{\omega}t + \phi_1) \quad \text{and} \quad x_2(t) = a_2 \sin(n\bar{\omega}t + \phi_2)$$

for $t \in [0, 2\pi/\bar{\omega}]$, where $\bar{\omega}, a_i, \phi_i$ are real numbers and n is an integer. Then, the closed curve in the shape space forms a Lissajous figure. Substituting the above $x_i(t)$ in (3.5), we have

$$\delta_z = \begin{cases} -\frac{\pi\epsilon^2(1-2\epsilon)la_1^2a_2m^2}{(I_z+2m\epsilon l^2)^2}\cos(\phi_2 - 2\phi_1) & \text{if } n = 2; \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $n = 2$ is the necessary condition for generating a nonzero geometric phase under the assumption (3.3). With this condition, $\phi_2 - 2\phi_1 = 2k\pi$, for $k = 0, \pm 1, \ldots$ gives the largest phase shift and $\phi_2 - 2\phi_1 = (2k+1)\pi/2$, for $k = 0, \pm 1, \ldots$, gives zero phase shift.

Next we formulate the optimal control problem for this particular mechanical system. For convenience, we re-scale the coordinate θ of S^1 by the factor $m\epsilon l/(I_z + 2m\epsilon l^2)^2$. Then the third equation of (3.1) becomes

$$\dot{\theta} = x_1^2\dot{x}_2 - x_2^2\dot{x}_1.$$

The optimal control problem is to find control $u_1(\cdot)$ and $u_2(\cdot)$ to

$$\text{minimize} \int_0^1 (u_1^2 + u_2^2)dt \tag{3.6a}$$

subject to

$$\begin{cases} \dot{x}_1 = u_1, \\ \dot{x}_2 = u_2, \\ \dot{\theta} = x_1^2u_2 - x_2^2u_1, \end{cases} \tag{3.6b}$$

with given boundary conditions

$$x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, \quad \theta(0) = \theta_0, \quad \theta(1) = \theta_1. \tag{3.6c}$$

Since the optimal control problem (3.6) has fixed boundary conditions, one needs to check the reachability for the system (3.6b). Define two vector fields on $\mathbb{R}^2 \times S^1$ by $g_1(q) = (1, 0, x_2)^T$ and $g_2(q) = (0, 1, -x_1^2)^T$. Then (3.6b) can be represented as

$$\dot{q} = g_1(q)u_1 + g_2(q)u_2.$$

It is easy to check that $[g_1, [g_1, g_2]](q) = (0, 0, 2)^T$ and $g_1(q), g_2(q), [g_1, [g_1, g_2]](q)$ are linearly independent for any $q = (x_1, x_2, \theta) \in \mathbb{R}^2 \times S^1$. By Chow's theorem, we conclude that, for the system (3.6b), there is an open set about the point $q = (0, 0, \theta)$ (or any other point in $\mathbb{R}^2 \times S^1$) such that any point in this set can be reached from q by a piecewise constant input (u_1, u_2).

Theorem 3.1 *If $(x_1(\cdot), x_2(\cdot), \theta(\cdot))$ is an optimal trajectory with control*

$$(u_1^*(\cdot), u_2^*(\cdot))$$

for the optimal control problem (3.6), *then there exists*

$$\lambda(\cdot) = (\lambda_1(\cdot), \lambda_2(\cdot), \lambda_3(\cdot))^T$$

on $[0,1]$ *satisfying the ordinary differential equations*

$$\begin{cases} \dot{x}_1 = u_1^*, \\ \dot{x}_2 = u_2^*, \\ \dot{\theta} = x_1^2 u_2^* - x_2^2 u_1^*, \end{cases} \qquad \begin{cases} \dot{\lambda}_1 = -2\lambda_3 x_1 u_2^*, \\ \dot{\lambda}_2 = 2\lambda_3 x_2 u_1^*, \\ \dot{\lambda}_3 = 0, \end{cases} \tag{3.7a}$$

where

$$u_1^*(q) = -\tfrac{1}{2}(\lambda_1 - \lambda_3 x_2^2) \quad \text{and} \quad u_2^*(q) = \tfrac{1}{2}(\lambda_2 + \lambda_3 x_1^2), \tag{3.7b}$$

with boundary conditions

$$x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, \quad \theta(0) = \theta_0, \quad \theta(1) = \theta_1.$$

Moreover, the system (3.7) *is completely integrable.*

Proof The equations (3.7) can be derived easily from variational principles. The derivation is omitted here. We just prove the solvability. From (3.7) one can get differential equations for the geodesics:

$$\begin{cases} \ddot{x}_1 - \lambda_3(x_1 + x_2)\dot{x}_2 = 0, \\ \ddot{x}_2 + \lambda_3(x_1 + x_2)\dot{x}_1 = 0, \\ \ddot{x}_3 + 2(x_2 - x_1)\dot{x}_1\dot{x}_2 + \lambda_3(x_1 + x_2)(x_1^2\dot{x}_1 + x_2^2\dot{x}_2) = 0, \end{cases} \tag{3.8}$$

for some constant λ_3. To integrate (3.8), let $w = x_1 + x_2$ and $v = x_1 - x_2$. We have

$$\begin{cases} \ddot{v} = \lambda_3 w\dot{w}, \\ \ddot{w} = -\lambda_3 w\dot{v}. \end{cases}$$

By integrating the first equation and substituting the result in the second equation, we get, for some constant c,

$$\ddot{w} + \lambda_3 w(c + \frac{\lambda_3}{2}w^2) = 0,$$

which is the equation for a quartic oscillator, solvable by elliptic function, i.e.,

$$t = \int \frac{dw}{\sqrt{C_1 + aw^2 + bw^4}} + C_2$$

for $a = \lambda_3 c/2$, $b = \lambda_3^2/8$, where C_1 and C_2 are integral constants. Therefore, we conclude that an optimal solution $(q(t), u(t))$ to (3.6) can be determined explicitly, i.e., the boundary value problem (3.6) is solvable. □

4 Three Dimensional System

Starting from this section, we assume that the system is free to move in three dimensional space. Again, we assume that the masses of the two oscillators are equal, i.e., $m_1 = m_2$. On the carrier a coordinate system 0-xyz is set such that the three axes are principal axes, i.e., the inertia tensor I_0 of the carrier can be represented as

$$I_0 = \text{diag}(I_x, I_y, I_z).$$

Two oscillators are allowed to move on the carrier such that Q_1 and Q_2 satisfy

$$Q_1(t) = (x_1(t)\cos(\psi_1),\ x_1(t)\sin(\psi_1),\ l)^T,$$
$$Q_2(t) = (x_2(t)\cos(\psi_2),\ x_2(t)\sin(\psi_2),\ -l)^T,$$

where $l > 0$ and ψ_i for $i = 1, 2$ are constants. This means that the two oscillators are restricted to move along their guide ways which are parallel to the 0-xy plane and are at an equal distance (l) from the plane (see Figure 4.1). The configuration space now becomes $Q = \mathbb{R}^2 \times SO(3)$ which will be coordinatized by (x_1, x_2, A), and the principal bundle is $(\mathbb{R}^2 \times SO(3), \mathbb{R}^2, SO(3))$.

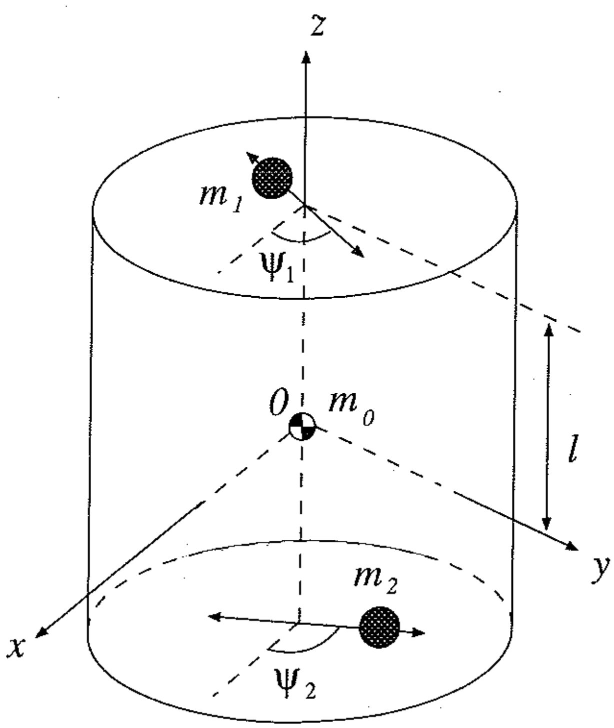

Figure 4.1 A Rigid Body with Two Oscillators: 3D System

By the above setting and condition $\mu = 0$, the angular velocity of the carrier in (2.14) is of the form

$$\Omega = \Omega_1(x_1, x_2)\dot{x}_1 + \Omega_2(x_1, x_2)\dot{x}_2, \tag{4.1}$$

where

$$\Omega_1 \triangleq \frac{1}{\det(I_{\text{lock}})}\begin{pmatrix}\omega_{11}\\ \omega_{12}\\ \omega_{13}\end{pmatrix} \quad \Omega_2 \triangleq \frac{1}{\det(I_{\text{lock}})}\begin{pmatrix}\omega_{21}\\ \omega_{22}\\ \omega_{23}\end{pmatrix}$$

with $\det(I_{\text{lock}})$ and ω_{ij} for $i = 1, 2$; $j = 1, 2, 3$ are polynomials of x_1 and x_2. And, the local connection form for the mechanical connection on $(\mathbb{R}^2 \times SO(3), \mathbb{R}^2, SO(3))$ is

$$\omega = -\Omega_1(x_1, x_2)\, dx_1 - \Omega_2(x_1, x_2)\, dx_2. \tag{4.2}$$

Although the above choice of parameters simplifies the bundle structure of the system, the expression of the connection form is still very complicated. In the rest of this section we will only consider some problems with special choices of ψ_i.

An interesting question is how to choose the constant parameters so that the above three dimensional system reduces to the planar system discussed in the preceding section. A natural guess may be that

$$\psi_1 = \psi_2 \triangleq \psi. \tag{4.3}$$

But this is not enough. In fact, when (4.3) is satisfied, ω_{13} and ω_{23} have simple expressions:

$$\omega_{13} = -\omega_{23} = \tfrac{1}{2}\epsilon^2(I_y - I_x)m^2l^2\sin(2\psi)(x_2 - x_1).$$

Thus, if $I_x = I_y$ or $\psi = 0$ or $\psi = \frac{\pi}{2}$, $\omega_{13} = \omega_{23} = 0$, i.e., the rigid body will only move (rotate) about the axis perpendicular to the plane formed by the guide ways of two oscillators. Otherwise, in general, the parallel motion of the two oscillators will cause the rigid body to drift about the z axis. In other words, when $\psi_1 = \psi_2$ holds, a sufficient condition for planar drift is that the carrier has axial symmetry about z-axis, or that the two oscillators move along the lines which are parallel with the same principal axis (0-x or 0-y) of the carrier.

Explicitly, if $\psi = 0$, i.e., both oscillators move parallel to principal axis 0-x, the local connection form is

$$\omega = (0,\ -\frac{\epsilon ml(dx_1 - dx_2)}{m\epsilon(\epsilon-1)x_2^2 + 2\epsilon^2 m x_1 x_2 + \epsilon m(\epsilon-1)x_1^2 - 2\epsilon ml^2 - I_y},\ 0)^T.$$

If $\psi = \frac{\pi}{2}$, i.e., both oscillators move parallel to principal axis 0-y, the local connection form is

$$\omega = (\frac{\epsilon ml(dx_1 - dx_2)}{m\epsilon(\epsilon-1)x_2^2 + 2\epsilon^2 m x_1 x_2 + \epsilon m(\epsilon-1)x_1^2 - 2\epsilon ml^2 - I_x},\ 0,\ 0)^T.$$

The nonzero terms in the above ω's are the same as ω in (3.2) (up to the choice of the coordinate system).

From the above discussion, it is apparent that if one is interested in full three dimensional motion of the carrier, some skewness in the directions of motion of the oscillators would be necessary. In the following, we set $\psi_1 = 0$ and $\psi_2 = \frac{\pi}{2}$ so that

$$Q_1(t) = (x_1(t),\ 0,\ -l)^T \quad \text{and} \quad Q_2(t) = (0,\ x_2(t),\ l)^T. \tag{4.4}$$

We will show that the kinematic control system corresponding to (1.9) with $\mu = 0$ is controllable in a neighborhood of $(0, 0, A)$ for any $A \in SO(3)$.

From (2.11) and (4.4) with $\mu = 0$, the kinematic control system is of the form

$$\begin{cases} \dot{x}_1 = u_1, \\ \dot{x}_2 = u_2, \\ \dot{A} = A(\widehat{\Omega}_1 u_1 + \widehat{\Omega}_2 u_2). \end{cases} \tag{4.5}$$

Let $x = (x_1, x_2)$ and $q = (x, A)$ be a point in Q. Equation (4.5) can be represented as

$$\dot{q} = X_1(q)u_1 + X_2(q)u_2, \tag{4.5$'$}$$

where $X_1(q) = ((1,0), A\widehat{\Omega}_1(x))$ and $X_2(q) = ((0,1), A\widehat{\Omega}_2(x))$.

Again, we can use the rank condition in Chow's theorem [4] to check the controllability of the system. To this end, we need a formula to compute the Lie bracket of vector fields X_1 and X_2 on $\mathbb{R}^2 \times SO(3)$, where X_i is represented as, in general,

$$X_i(x, A) = (F_i(x), A\widehat{G}_i(x)) \qquad i = 1, 2 \tag{4.6}$$

for smooth mappings $F_i : \mathbb{R}^2 \to \mathbb{R}^2$ and $G_i : \mathbb{R}^2 \to \mathbb{R}^3$.

Proposition 4.1 *Given two vector fields X_1 and X_2 on $\mathbb{R}^2 \times SO(3)$ shown in (4.6), the Lie bracket of X_1 and X_2 is given by*

$$[X_1, X_2](x, A) = \Big(\frac{\partial F_2}{\partial x} F_1 - \frac{\partial F_1}{\partial x} F_2,\ A\Big[G_1 \times G_2 + \frac{\partial G_2}{\partial x} F_1 - \frac{\partial G_1}{\partial x} F_2\Big]^{\widehat{}}\ \Big) \tag{4.7}$$

for any point $(x, A) \in \mathbb{R}^2 \times SO(3)$.

Proof Let

$$\phi_i(\tau) = (x + \tau F_i(x), A\exp(\tau \widehat{G}_i(x)))$$

be an integral curve of X_i at (x, A) for $i = 1, 2$. Then,

$$\begin{aligned}
[X_1, X_2](x, A) = \frac{d}{d\tau}\Big|_{\tau=0} & (F_2(x + \tau F_1) - F_1(x + \tau F_2), \\
A\exp(\tau\widehat{G}_1)\widehat{G}_2(x+\tau F_1) &- A\exp(\tau \widehat{G}_2)\widehat{G}_1(x + \tau F_2)) \\
&= \Big(\frac{\partial F_2}{\partial x} F_1 - \frac{\partial F_1}{\partial x} F_2,\ A\Big(\widehat{G}_1\widehat{G}_2 - \widehat{G}_2\widehat{G}_1 + \Big[\frac{\partial G_2}{\partial x} F_1 - \frac{\partial G_1}{\partial x} F_2\Big]^{\widehat{}}\Big)\Big) \\
&= \Big(\frac{\partial F_2}{\partial x} F_1 - \frac{\partial F_1}{\partial x} F_2,\ A\Big[G_1 \times G_2 + \frac{\partial G_2}{\partial x} F_1 - \frac{\partial G_1}{\partial x} F_2\Big]^{\widehat{}}\Big).
\end{aligned}$$

□

Theorem 4.2 *For system (4.5), there is an open set about $(0,0,A)$ for any $A \in SO(3)$ such that any point in this set can be reached from $(0,0,A)$ by a piecewise constant input (u_1, u_2).*

Proof Let

$$X_3(x,A) \triangleq [X_1, X_2](x,A) = (F_3(x), A\widehat{G}_3(x)),$$
$$X_4(x,A) \triangleq [X_1, X_3](x,A) = (F_4(x), A\widehat{G}_4(x)),$$
$$X_5(x,A) \triangleq [X_2, X_3](x,A) = (F_5(x), A\widehat{G}_5(x)),$$

where F_i and G_i, $i = 1,2,3$ are computed by using (4.7) and (4.5)′. It is easy to see that

$$F_i = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{for } i = 3,4,5.$$

Using Macsyma, one can check that, for Q_1 and Q_2 given in (4.4),

$$\det(G_3, G_4, G_5)|_{x=(0,0)} \neq 0.$$

Since $T_{(x,A)}(\mathbb{R}^2 \times SO(3)) \simeq \mathbb{R}^5$ and the vector fields shown in (4.5)′. which generate the smallest involutive distribution, have special forms, namely $F_1(x) = (1,0)^T$ and $F_2(x) = (0,1)^T$, the above equations are sufficient to show that vector fields X_i for $i = 1, ..., 5$ are independent. Consequently, the control system given in (4.5) is controllable by Chow's theorem. □

We now turn to the optimal control problem. Corresponding to (P2) in Section 1, the goal here is to find $u_1(\cdot)$ and $u_2(\cdot)$ to

$$\text{minimize} \int_0^1 (u_1^2 + u_2^2)\, dt \tag{4.8a}$$

subject to

$$\begin{cases} \dot{x}_1 = u_1, \\ \dot{x}_2 = u_2, \\ \dot{A} = A(\widehat{\Omega}_1 u_1 + \widehat{\Omega}_2 u_2), \end{cases} \tag{4.8b}$$

for given boundary conditions

$$\begin{aligned} &x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, \\ &A(0) = A_0 \in SO(3), \\ &A(1) = A_1 \in SO(3). \end{aligned} \tag{4.8c}$$

The necessary conditions for the above problem are given in the following theorem.

Theorem 4.3 *If for the optimal control problem given by (4.8), $(x(\cdot), A(\cdot))$ is an optimal trajectory with controls $(u_1^*(\cdot), u_2^*(\cdot))$ then there exist*

$$\mu(\cdot) = (\mu_1(\cdot), \mu_2(\cdot)) \quad \textit{and} \quad \lambda(\cdot) \textit{ on } [0,1]$$

satisfying the ordinary differential equations

$$\begin{cases} \dot{x}_1 = u_1^*, \\ \dot{x}_2 = u_2^*, \\ \dot{A} = A(\widehat{\Omega}_1 u_1^* + \widehat{\Omega}_2 u_2^*), \end{cases} \qquad \begin{cases} \dot{\mu}_1 = -\lambda^T(\frac{\partial \Omega_1}{\partial x_1} u_1^* + \frac{\partial \Omega_2}{\partial x_1} u_2^*), \\ \dot{\mu}_2 = -\lambda^T(\frac{\partial \Omega_1}{\partial x_2} u_1^* + \frac{\partial \Omega_2}{\partial x_2} u_2^*), \\ \dot{\lambda} = \lambda \times (\Omega_1 u_1^* + \Omega_2 u_2^*), \end{cases} \tag{4.9a}$$

where

$$\begin{aligned} u_1^*(x, A, \mu, \lambda) &= -\tfrac{1}{2}(\mu_1 + \lambda^T \Omega_1), \\ u_2^*(x, A, \mu, \lambda) &= -\tfrac{1}{2}(\mu_2 + \lambda^T \Omega_2), \end{aligned} \tag{4.9b}$$

with boundary conditions

$$\begin{gathered} x_1(0) = x_1(1) = x_2(0) = x_2(1) = 0, \\ A(0) = A_0 \in SO(3), \quad A(1) = A_1 \in SO(3). \end{gathered} \tag{4.9c}$$

Proof Here, we first consider a slightly more general form of the problem (4.8). The optimal control problem now is to determine a control $(u_1(\cdot), u_2(\cdot))$ to

$$\text{minimize} \int_0^1 (u_1^2 + u_2^2)\, dt \tag{4.10a}$$

subject to

$$\begin{aligned} \dot{x} &= f_1(x)u_1 + f_2(x)u_2, \\ \dot{A} &= A(\widehat{\Omega}_1(x)u_1 + \widehat{\Omega}_2(x)u_2), \end{aligned} \tag{4.10b}$$

with boundary conditions

$$(x(0), A(0)) = (x_0, A_0) \quad \text{and} \quad (x(1), A(1)) = (x_1, A_1), \tag{4.10c}$$

where $(x(t), A(t)) \in \mathbb{R}^2 \times SO(3), \forall t \in [0,1]$, $f_i : \mathbb{R}^2 \to \mathbb{R}^2$, $\Omega_i : \mathbb{R}^2 \to \mathbb{R}^3$ for $i = 1, 2$ and $u(\cdot) = (u_1(\cdot), u_2(\cdot)) : [0,1] \to \mathbb{R}^2$ is a piecewise smooth function. Applying the maximum principle to this problem, we have the following result.

Denoting by $z = ((x, A), (\mu, \Lambda_A)) = ((x, A), (\mu, A\widehat{\lambda})) \simeq ((x, A), (\mu, \lambda))$ a point in $T^*(\mathbb{R}^2 \times SO(3))$, we define *the pseudo-Hamiltonian,*

$$\mathcal{H} : T^*(\mathbb{R}^2 \times SO(3)) \times \mathbb{R}^2 \to \mathbb{R},$$

by

$$\begin{aligned} \mathcal{H}(z, u) &\triangleq \langle u, u \rangle + \langle \mu, f_1(x)u_1 + f_2(x)u_2 \rangle + \langle \Lambda_A, A(\widehat{\Omega}_1(x)u_1 + \widehat{\Omega}_2(x)u_2) \rangle \\ &= \langle u, u \rangle + \langle \mu, f_1(x)u_1 + f_2(x)u_2 \rangle + \langle \lambda, \Omega_1(x)u_1 + \Omega_2 u_2 \rangle. \end{aligned}$$

Here, we have used $\langle \cdot, \cdot \rangle$ to denote the real-valued pairing of a vector space and its dual. Define Hamiltonian

$$H(z) \triangleq \min_{u \in \mathbb{R}^2} \mathcal{H}(z, u).$$

Since the control space is unbounded, one can find functions $u^*(z) = (u_1^*(z), u_2^*(z))$ on $T^*(\mathbb{R}^2 \times SO(3))$ with

$$\begin{aligned} u_1^*(z) &= -\tfrac{1}{2}(\mu^T f_1(x) + \lambda^T \Omega_1(x)), \\ u_2^*(z) &= -\tfrac{1}{2}(\mu^T f_2(x) + \lambda^T \Omega_2(x)), \end{aligned} \tag{4.11}$$

so that

$$H(z) = \mathcal{H}(z, u^*(z)) = -\tfrac{1}{4}\sum_{i=1}^{2}(\langle \mu, f_i \rangle + \langle \lambda, \Omega_i \rangle)^2. \tag{4.12}$$

From the maximum principle, we know that the existence of an optimal control for problem (4.10) implies that there is a solution to the following system on $T^*(\mathbb{R}^2 \times SO(3))$:

$$\dot{z} = X_H(z) \tag{4.13}$$

with

$$\tau(z(0)) = (x_0, A_0) \quad \text{and} \quad \tau(z(1)) = (x_1, A_1),$$

where X_H is the Hamiltonian vector field with respect to the Hamiltonian (4.12) and $\tau : T^*(\mathbb{R}^2 \times SO(3)) \to \mathbb{R}^2 \times SO(3)$ is the canonical projection. Our goal now is to determine the vector field X_H on $T^*(\mathbb{R}^2 \times SO(3))$.

Recall that, for given n-dimensional smooth manifold Q, its cotangent bundle T^*Q has a canonical symplectic form Ω_0 ([10]). Given a Hamiltonian H on T^*Q, the corresponding Hamiltonian vector field X_H on T^*Q is defined by

$$\Omega_0(X_H, Y) = H(Y) \tag{4.14}$$

for any vector field $Y \in \mathfrak{X}(T^*Q)$. The local expression of (4.14) can be given as (cf. Theorem 3.2.10 in [10])

$$\omega(x, \alpha)((x, \alpha, e_1, \beta_1), (x, \alpha, e_2, \beta_2)) = \langle \beta_2, e_1 \rangle - \langle \beta_1, e_2 \rangle \tag{4.15}$$

for $(x, \alpha) \in T^*Q$ and $(e_i, \beta_i) \in T_{(x,\alpha)}T^*Q$, $i = 1, 2$. In our problem, $Q = \mathbb{R}^2 \times SO(3)$. For the Hamiltonian given in (4.12), the corresponding Hamiltonian vector field X_H in (4.13) will be determined by (4.14).

Let

$$y(t) = ((x + tv, Ae^{t\widehat{\alpha}}), (\mu + t\phi, Ae^{t\widehat{\alpha}}(\widehat{\lambda} + t\widehat{\beta})))$$

be an integral curve on $T^*(\mathbb{R}^2 \times SO(3))$ at

$$z = ((x, A), (\mu, A\widehat{\lambda})) \qquad \text{for any } v, \phi \in \mathbb{R}^2,\ \alpha, \beta \in \mathbb{R}^3.$$

Then its tangent vector at z is given by

$$Y(z) = ((x, A), (\mu, A\widehat{\lambda}), (v, A\widehat{\alpha}), (\phi, A(\widehat{\widehat{\alpha}\lambda} + \widehat{\beta}))).$$

Now the right-hand-side of (4.14) can be calculated as

$$
\begin{aligned}
H(Y) &= \frac{d}{dt}\bigg|_{t=0} H(y(t)) \\
&= -\tfrac{1}{4}\frac{d}{dt}\bigg|_{t=0} \sum_{i=1}^{2}(\langle \mu + t\phi, f_i(x+tv)\rangle + \langle \lambda + t\beta, \Omega_i(x+tv)\rangle)^2 \\
&= \langle \phi, f_1 u_1^* + f_2 u_2^*\rangle + \langle \beta, \Omega_1 u_1^* + \Omega_2 u_2^*\rangle \\
&\quad + \langle \mu, (Df_1 u_1^* + Df_2 u_2^*)\cdot v\rangle + \langle \lambda, (D\Omega_1 u_1^* + D\Omega_2 u_2^*)\cdot v\rangle, \qquad (4.16)
\end{aligned}
$$

where u_i^* is given in (4.11) and

$$
Df_i = \begin{pmatrix} \frac{\partial f_i}{\partial x_1} & \frac{\partial f_i}{\partial x_2} \end{pmatrix} \quad \text{and} \quad D\Omega_i = \begin{pmatrix} \frac{\partial \Omega_i}{\partial x_1} & \frac{\partial \Omega_i}{\partial x_2} \end{pmatrix}
$$

for $i = 1, 2$ are 2×2 and 3×2 matrices, respectively.

On the other hand, let

$$
X_H(z) = ((x, A), (\mu, A\widehat{\lambda}), (w, A\widehat{\epsilon}), (\eta, A(\widehat{\epsilon\lambda} + \widehat{\delta})))
$$

for some vectors $w, \eta \in \mathbb{R}^2$ and $\epsilon, \delta \in \mathbb{R}^3$ which will be determined. Applying (4.15), we have

$$
\begin{aligned}
\Omega_0(z)(X_H, Y) &= \langle (\phi, A(\widehat{\alpha\lambda} + \widehat{\beta})), (w, A\widehat{\epsilon})\rangle - \langle (\eta, A(\widehat{\epsilon\lambda} + \widehat{\delta})), (v, A\widehat{\alpha})\rangle \\
&= \langle \phi, w\rangle - \langle \eta, v\rangle + \langle \beta, \epsilon\rangle - \tfrac{1}{2}\,\mathrm{tr}([\widehat{\epsilon}, \widehat{\lambda}] + \widehat{\delta})\widehat{\alpha}. \qquad (4.17)
\end{aligned}
$$

In order to have equation (4.14), we need to make (4.16) equal to (4.17). This leads to following choice of w, ϵ, η and δ.

$$
\begin{aligned}
w &= f_1 u_1^* + f_2 u_2^*, \\
\eta &= -(Df_1 u_1^* + Df_2 u_2^*)^T \mu - (D\Omega_1 u_1^* + D\Omega_2 u_2^*)^T \lambda, \\
\epsilon &= \Omega_1 u_1^* + \Omega_2 u_2^*, \\
\delta &= \lambda \times (\Omega_1 u_1^* + \Omega_2 u_2^*).
\end{aligned} \qquad (4.18)
$$

With the above equations, the vector field X_H can be completely determined.

In summary, the differential equation (4.13) now has the following form.

$$
\begin{aligned}
\dot{x} &= f_1(x)u_1^* + f_2(x)u_2^*, \\
\dot{A} &= A(\widehat{\Omega}_1(x)u_1^* + \widehat{\Omega}_2(x)u_2^*), \\
\dot{\mu} &= -(Df_1(x)u_1^* + Df_2(x)u_2^*)^T \mu - (D\Omega_1(x)u_1^* + D\Omega_2(x)u_2^*)^T \lambda, \\
\dot{\lambda} &= \lambda \times (\Omega_1(x)u_1^* + \Omega_2(x)u_2^*),
\end{aligned} \qquad (4.19a)
$$

where

$$
\begin{aligned}
u_1^* &= -\tfrac{1}{2}(\mu^T f_1(x) + \lambda^T \Omega_1(x)), \\
u_2^* &= -\tfrac{1}{2}(\mu^T f_2(x) + \lambda^T \Omega_2(x)).
\end{aligned} \qquad (4.19b)
$$

When $f_1 = (1, 0)^T$ and $f_2 = (0, 1)^T$, (4.19) leads to (4.9). □

Remark 4.4 From the proof of Theorem 4.3, we can see that the geometric treatment of the optimal control problem allows us to define a Hamiltonian vector field on the manifold $T^*(\mathbb{R}^2 \times SO(3))$. The solution of the optimal control problem will correspond to a trajectory of this vector field. Of course, in general, it is almost impossible to find explicit solutions, although we did find one for the planar system as shown in Section 3. However, as we will show in the next section, identification of symmetries in such a Hamiltonian system will allow a reduction in the order of the system by applying symplectic or Poisson reduction theory [11, 12]. □

5 Symmetry and Reduction

Recall that the manifold $P = T^*(\mathbb{R}^2 \times SO(3))$, parameterized by

$$z = (x, A, \mu, A\widehat{\lambda}),$$

is symplectic. Hence, it is also Poisson. One can verify that the Poisson bracket of functions F_1 and F_2 on P is given by

$$\begin{aligned}\{F_1, F_2\}_P(z) =& \frac{\partial F_1}{\partial x} \cdot \frac{\partial F_2}{\partial \mu} - \frac{\partial F_2}{\partial x} \cdot \frac{\partial F_1}{\partial \mu} \\ &+ \langle D_A F_1, D_{A\widehat{\lambda}} F_2 \rangle - \langle D_A F_2, D_{A\widehat{\lambda}} F_1 \rangle.\end{aligned} \tag{5.1}$$

In the preceding section we have shown that, for the Hamiltonian

$$H(z) = -\tfrac{1}{4}\sum_{i=1}^{2}(\mu_i + \langle \lambda, \Omega_i \rangle)^2, \tag{5.2}$$

the Hamiltonian vector field is

$$X_H(z) = \begin{pmatrix} u_1^* \\ u_2^* \\ A(\widehat{\Omega}_1 u_1^* + \widehat{\Omega}_2 u_2^*) \\ -\lambda^T\left(\frac{\partial \Omega_1}{\partial x_1} u_1^* + \frac{\partial \Omega_2}{\partial x_1} u_2^*\right) \\ -\lambda^T\left(\frac{\partial \Omega_1}{\partial x_2} u_1^* + \frac{\partial \Omega_2}{\partial x_2} u_2^*\right) \\ \lambda \times (\Omega_1 u_1^* + \Omega_2 u_2^*) \end{pmatrix}, \tag{5.3}$$

where

$$u_1^*(z) = -\tfrac{1}{2}(\mu_1 + \lambda^T \Omega_1), \quad u_2^*(z) = -\tfrac{1}{2}(\mu_2 + \lambda^T \Omega_2).$$

Consider an action of $SO(3)$ on $\mathbb{R}^2 \times SO(3)$ given by

$$\begin{aligned}\Phi : SO(3) \times (\mathbb{R}^2 \times SO(3)) &\to \mathbb{R}^2 \times SO(3), \\ (B, (x, A)) &\mapsto (x, BA),\end{aligned}$$

and its cotangent lift on $T^*(\mathbb{R}^2 \times SO(3))$,

$$\begin{aligned}\Phi^{T^*} : SO(3) \times T^*(\mathbb{R}^2 \times SO(3)) &\to T^*(\mathbb{R}^2 \times SO(3)), \\ (B, (x, A, \mu, A\widehat{\lambda})) &\mapsto (x, BA, \mu, BA\widehat{\lambda}).\end{aligned}$$

The quotient space $\widetilde{P} = T^*(\mathbb{R}^2 \times SO(3))/SO(3) \simeq \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^3$ can be coordinatized by $\widetilde{z} = (x, \mu, \lambda)$. Let $\widetilde{\pi}$ be the canonical projection from P to $\widetilde{P}$. We then have the following theorem.

Theorem 5.1 *The Hamiltonian system $(P, \{\ ,\ \}_P, X_H)$ has $SO(3)$ symmetry and is Poisson reducible. The Poisson reduced system is given by*

$$\dot{\widetilde{z}} = \widetilde{\Lambda}(\widetilde{z}) \nabla \widetilde{H}, \tag{5.4}$$

where $\widetilde{\Lambda}(\widetilde{z})$ is the Poisson structure given by

$$\widetilde{\Lambda}(\widetilde{z}) = \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & \widehat{\lambda} \end{pmatrix} \tag{5.5}$$

for I denoting the 2×2 identity matrix and 0's null matrices with suitable dimensions, and $\widetilde{H}$ the reduced Hamiltonian given by $\widetilde{H} \circ \widetilde{\pi} = H$. In addition, the Casimir functions of the Poisson reduced system are all real-valued smooth functions of $\|\lambda\|^2$ which is a first integral for the system, i.e., for some constant C_1,

$$\|\lambda\|^2 = C_1. \tag{5.6}$$

Proof It is obvious that the Hamiltonian (5.2) is invariant under the action Φ^{T^*} since it does not have A in its expression. This immediately implies the $SO(3)$-symmetry for the system. Moreover, the reduced Hamiltonian $\widetilde{H}$ on $T^*(\mathbb{R}^2 \times SO(3))/SO(3)$ is simply

$$\widetilde{H}(x, \mu, \lambda) = H(x, A, \mu, A\widehat{\lambda}). \tag{5.7}$$

Let f_1 and f_2 be smooth functions on $\widetilde{P} = T^*(\mathbb{R}^2 \times SO(3))/SO(3)$. Let F_1 and F_2 be lifted functions on $P = T^*(\mathbb{R}^2 \times SO(3))$ such that

$$F_i(x, A, \mu, A\widehat{\lambda}) = f_i(x, \mu, \lambda), \qquad i = 1, 2.$$

We need to find the expression of $\{f_1, f_2\}_{\widetilde{P}}$ such that

$$\{f_1, f_2\}_{\widetilde{P}}(x, \mu, \lambda) = \{F_1, F_2\}_P(x, A, \mu, A\widehat{\lambda}), \tag{5.8}$$

where $\{F_1, F_2\}_P$ is given in (5.1). As we have seen in the proof of Theorem 4.3, a tangent vector Y on $T^*(\mathbb{R}^2 \times SO(3))$ at $z = (x, A, \mu, A\widehat{\lambda})$ has the form

$$Y(z) = (v, A\widehat{\alpha}, w, A(\widehat{\alpha}\widehat{\lambda} + \widehat{\beta}))_z,$$

where $v, w \in \mathbb{R}^2$ and $\alpha, \beta \in \mathbb{R}^3$. An integral curve of Y at z can be written as,

$$y(t) = (x + tv, Ae^{t\widehat{\alpha}}, \mu + tw, Ae^{t\widehat{\alpha}}(\widehat{\lambda} + t\widehat{\beta})).$$

Then by the definition of F_i we have

$$\begin{aligned} dF_i \cdot Y(z) &= \frac{d}{dt}|_{t=0} F_i(x+tv, Ae^{t\widehat{\alpha}}, \mu + tw, Ae^{t\widehat{\alpha}}(\widehat{\lambda} + t\widehat{\beta})) \\ &= \frac{d}{dt}|_{t=0} f_i(x+tv, \mu+tw, \lambda + t\beta) \\ &= \frac{\partial f_i}{\partial x} \cdot v + \frac{\partial f_i}{\partial \mu} \cdot w + \frac{\partial f_i}{\partial \lambda} \cdot \beta. \end{aligned} \tag{5.9}$$

On the other hand

$$\begin{aligned} dF_i \cdot Y(z) &= \frac{\partial F_i}{\partial x} \cdot v + \langle D_A F_i, A\alpha\rangle + \frac{\partial F_i}{\partial \mu} \cdot w + \langle D_{A\widehat{\lambda}} F_i, A(\widehat{\alpha}\widehat{\lambda} + \beta)\rangle \\ &= \frac{\partial F_i}{\partial x} \cdot v + \frac{\partial F_i}{\partial \mu} \cdot w + \langle D_A F_i - (D_{A\widehat{\lambda}} F_i)\widehat{\lambda}, A\widehat{\alpha}\rangle + \langle D_{A\widehat{\lambda}} F_i, A\widehat{\beta}\rangle. \end{aligned} \tag{5.10}$$

Comparing (5.9) with (5.10), we get

$$\frac{\partial F_i}{\partial x} = \frac{\partial f_i}{\partial x}, \quad \frac{\partial F_i}{\partial \mu} = \frac{\partial f_i}{\partial \mu}, \quad D_{A\widehat{\lambda}} F_i = A\widehat{\frac{\partial f_i}{\partial \lambda}}, \quad D_A F_i = A\widehat{\frac{\partial f_i}{\partial \lambda}}\widehat{\lambda}. \tag{5.11}$$

From (5.1), (5.8) and (5.11), we have

$$\begin{aligned} \{f_1, f_2\}_{\widetilde{P}}(\widetilde{z}) &= \left\langle A\widehat{\frac{\partial f_1}{\partial \lambda}}\widehat{\lambda}, A\widehat{\frac{\partial f_2}{\partial \lambda}}\right\rangle - \left\langle A\widehat{\frac{\partial f_2}{\partial \lambda}}\widehat{\lambda}, A\widehat{\frac{\partial f_1}{\partial \lambda}}\right\rangle \\ &\quad + \frac{\partial f_1}{\partial x} \cdot \frac{\partial f_2}{\partial \mu} - \frac{\partial f_2}{\partial x} \cdot \frac{\partial f_1}{\partial \mu} \\ &= -\left(\frac{\partial f_1}{\partial \lambda} \times \frac{\partial f_2}{\partial \lambda}\right) \cdot \lambda + \frac{\partial f_1}{\partial x} \cdot \frac{\partial f_2}{\partial \mu} - \frac{\partial f_2}{\partial x} \cdot \frac{\partial f_1}{\partial \mu} \\ &= \begin{pmatrix} \frac{\partial f_1}{\partial x}^T & \frac{\partial f_1}{\partial \mu}^T & \frac{\partial f_1}{\partial \lambda}^T \end{pmatrix} \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & \widehat{\lambda} \end{pmatrix} \begin{pmatrix} \frac{\partial f_2}{\partial x} \\ \frac{\partial f_2}{\partial \mu} \\ \frac{\partial f_2}{\partial \lambda} \end{pmatrix}, \end{aligned}$$

where I is 2×2 identity matrix. Therefore, the matrix in (5.5) is the reduced Poisson structure. The proof of the rest of this Theorem can be easily carried out. □

Remark 5.2 Since we are dealing with a trivial principal bundle here with structure group as the symmetry group, after recognizing the coordinates for $\widetilde{P}$ one should be able to find the Poisson reduced system by eliminating the third equation of (4.9a) and determine the reduced Poisson structure from it. The first integral in (5.6) follows also from the last equation of (4.9a). □

6 Approximation and Further Reduction

Up to now, we have made no simplifications or approximations of the optimal control problem. It is customary to make *ad hoc* approximations in the interest of ensuring analytic integrability or numerical solvability. However, in the process one can easily destroy symmetries inherent to the problem. This is highly undesirable. On the other hand, in many physical systems, simplification and/or approximation may bring symmetries to the system. Our purpose in this section is to impose suitable assumptions, explore further symmetry and reduce the order of the system again.

As in Section 3, we assume that the distances of the point masses to 0-z axis, x_i, are very small in comparison with their distances to the xy plane, i.e.,

$$|x_i|/l \ll 1. \tag{6.1}$$

By doing so, we ignore the higher order terms with

$$\left(\frac{x_1}{l}\right)^i \left(\frac{x_2}{l}\right)^j \qquad \text{for } i+j>2$$

in both numerators and denominators of $\Omega_1(x_1, x_2)$ and $\Omega_2(x_1, x_2)$. In addition, inspired by the symmetric heavy top, we posit one more important assumption. We assume that the rigid body is symmetric about 0-z axis which implies

$$I_x = I_y. \tag{6.2}$$

Under the above approximation and assumption, the angular velocity Ω takes the form:

$$\Omega = \begin{pmatrix} 0 \\ a \\ bx_2 \end{pmatrix} \dot{x}_1 + \begin{pmatrix} a \\ 0 \\ -bx_1 \end{pmatrix} \dot{x}_2, \tag{6.3}$$

where

$$\begin{aligned} a &= -\frac{1}{\Delta}\epsilon m I_z l(2\epsilon m l^2 + I_x), \\ b &= \frac{1}{\Delta}\epsilon^2 m(2\epsilon m l^2 + I_x)(2\epsilon m l^2 - m l^2 + I_x), \\ \Delta &= I_z(2\epsilon m l^2 + I_x)^2. \end{aligned}$$

The approximation of $\widetilde{H}$ is, for $\widetilde{z} = (x, \mu, \lambda)$,

$$H_{12}(\widetilde{z}) = -\tfrac{1}{4}((b\lambda_3 x_2 + \mu_1 + a\lambda_2)^2 + (b\lambda_3 x_1 - \mu_2 - a\lambda_1)^2). \tag{6.4}$$

Applying H_{12} to (5.4), we get corresponding Poisson dynamics

$$\dot{\widetilde{z}} = \widetilde{\Lambda}(\widetilde{z})\nabla H_{12}$$

or, explicitly,

$$\begin{cases} \dot{x}_1 = -\dfrac{b\lambda_3 x_2 + \mu_1 + a\lambda_2}{2}, \\ \dot{x}_2 = \dfrac{b\lambda_3 x_1 - \mu_2 - a\lambda_1}{2}, \\ \dot{\mu}_1 = \dfrac{b^2\lambda_3^2 x_1 - b\lambda_3\mu_2 - ab\lambda_1\lambda_3}{2}, \\ \dot{\mu}_2 = \dfrac{b^2\lambda_3^2 x_2 + b\lambda_3\mu_1 + ab\lambda_2\lambda_3}{2}, \\ \dot{\lambda}_1 = -\tfrac{1}{2}(b^2\lambda_2\lambda_3 x_2^2 + (b\lambda_2\mu_1 - ab\lambda_3^2 + ab\lambda_2^2)x_2 + b^2\lambda_2\lambda_3 x_1^2 \\ \qquad + (-b\lambda_2\mu_2 - ab\lambda_1\lambda_2)x_1 - a\lambda_3\mu_1 - a^2\lambda_2\lambda_3), \\ \dot{\lambda}_2 = \tfrac{1}{2}(b^2\lambda_1\lambda_3 x_2^2 + (b\lambda_1\mu_1 + ab\lambda_1\lambda_2)x_2 + b^2\lambda_1\lambda_3 x_1^2 \\ \qquad + (-b\lambda_1\mu_2 + ab\lambda_3^2 - ab\lambda_1^2)x_1 - a\lambda_3\mu_2 - a^2\lambda_1\lambda_3), \\ \dot{\lambda}_3 = -\dfrac{ab\lambda_1\lambda_3 x_2 + ab\lambda_2\lambda_3 x_1 - a\lambda_2\mu_2 + a\lambda_1\mu_1}{2}. \end{cases} \tag{6.5}$$

Next, we show that the above system admits a symmetry group and the order of the reduced Hamiltonian system can be finally brought down to 4. Consider a one-parameter group $G_0 \simeq S^1$ with each element having the form:

$$g_\tau = \mathrm{Diag}(\mathrm{Rot}(\tau), \quad \mathrm{Rot}(\tau), \mathrm{Rot}^3(-\tau)), \tag{6.6}$$

where

$$\mathrm{Rot}(\tau) = \begin{pmatrix} \cos(\tau) & \sin(\tau) \\ -\sin(\tau) & \cos(\tau) \end{pmatrix} \quad \text{and} \quad \mathrm{Rot}^3(\tau) = \begin{pmatrix} \cos(\tau) & \sin(\tau) & 0 \\ -\sin(\tau) & \cos(\tau) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Define the action Ψ of G_0 on $\widetilde{P} = T^*(\mathbb{R}^2 \times SO(3))/SO(3) \simeq \mathbb{R}^7$ by

$$\begin{aligned} &\Psi : G_0 \times \mathbb{R}^7 \to \mathbb{R}^7, \\ &(\tau, (x, \mu, \lambda)) \mapsto (\mathrm{Rot}(\tau)x, \mathrm{Rot}(\tau)\mu, \mathrm{Rot}^3(-\tau)\lambda). \end{aligned} \tag{6.7}$$

We then have the following striking fact.

Theorem 6.1 *Following the rule of* (6.7), *the group G_0 acts on the Poisson manifold $\widetilde{P}$ canonically, i.e., for any $g_\tau \in G_0$,*

$$\{f_1, f_2\}_{\widetilde{P}} \circ \Psi_{g_\tau} = \{f_1 \circ \Psi_{g_\tau}, f_2 \circ \Psi_{g_\tau}\}_{\widetilde{P}}, \qquad \forall f_i \in C^\infty(\widetilde{P}), \forall g_\tau \in G_0.$$

In addition, the approximate Hamiltonian H_{12} is G_0-invariant, i.e.,

$$H_{12}(\widetilde{z}) = H_{12}(\Psi_{g_\tau}(\widetilde{z})).$$

Proof The first assertion is equivalent to

$$D\Psi_{g_\tau}(\tilde{z})\tilde{\Lambda}(\tilde{z})D\Psi_{g_\tau}(\tilde{z})^T = \tilde{\Lambda}(\Psi_{g_\tau}(\tilde{z})).$$

This can be shown by a straightforward calculation. So is the second assertion.

□

Remark 6.2 From this theorem, one immediately concludes that G_0 is a symmetry group of the system (6.5). □

Recall that, given a Poisson manifold $(M, \{\ ,\ \}_M)$ and a function H on M, the Hamiltonian vector field X_H is defined by

$$X_H[F] = \{F, H\}_M \qquad \forall F \in C^\infty(M). \tag{6.8}$$

Let the group G act canonically on M by the action $\Phi : G \times M \to M$. A momentum map $\mathbf{J} : M \to \mathcal{G}^*$ (the dual of Lie algebra of G) of this action is defined by

$$\langle \mathbf{J}(z), \xi \rangle = J(\xi)(z) \tag{6.9}$$

for all $\xi \in \mathcal{G}$ and $z \in M$, where $J : \mathcal{G} \to C^\infty(M)$ is a linear map such that

$$X_{J(\xi)} = \xi_M. \tag{6.10}$$

From (6.8)-(6.10), we see that the momentum map can be determined by the following equation:

$$\{F, J(\xi)\}_M = \xi_M[F]. \tag{6.11}$$

The Hamiltonian version of Noether's theorem states that if the Lie group, G, acting canonically on the Poisson manifold M admits a momentum mapping $\mathbf{J} : M \to \mathcal{G}^*$ and $H \in C^\infty(M)$ is G-invariant, i.e.,

$$H \circ \Phi_g = H \quad \text{or} \quad \xi_M[H] = 0 \ \text{ for all } \xi \in \mathcal{G},$$

then $\mathbf{J}$ is a constant of the motion for H, i.e.,

$$\mathbf{J} \circ \phi_t = \mathbf{J},$$

where ϕ_t is the flow of X_H.

We now return to our problem. It is clear that by Theorem 6.1 conditions of Noether's theorem are satisfied and, consequently, the admitted momentum map will be an integral of the reduced system (6.5). Indeed, we have the following theorem.

Theorem 6.3 *The constant momentum map for the system $(\tilde{P}, \{\ ,\ \}_{\tilde{P}})$ corresponding to the action Ψ of G_0 on $\tilde{P}$ is*

$$J(\xi)(\tilde{z}) = -x_1\mu_2 + x_2\mu_1 + \lambda_3. \tag{6.12}$$

Proof We will determine a function $J(\xi)$ which satisfies

$$\{f, J(\xi)\}_{\widetilde{P}} = \xi_{\widetilde{P}}(f), \qquad \forall f \in C^\infty(\widetilde{P}),\ \forall \xi \in \mathcal{G}_0.$$

From (6.7) we know that $\xi \in \mathcal{G}_0$ is of the form

$$\xi = \xi' \,\mathrm{Diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right),$$

where ξ' is any constant in $\mathbb{R}$ which will later be chosen to be 1. It is easy to show that the infinitesimal generator of ξ is given by

$$\xi_{\widetilde{P}}(\widetilde{z}) = \frac{d}{d\tau}\bigg|_{\tau=0} \Psi(\exp(\xi\tau), \widetilde{z}) = (x_2, -x_1, \mu_2, -\mu_1, -\lambda_2, \lambda_1, 0)^T.$$

Then, for any smooth function f on $\widetilde{P}$,

$$\xi_{\widetilde{P}}(f)(\widetilde{z}) = x_2\frac{\partial f}{\partial x_1} - x_1\frac{\partial f}{\partial x_2} + \mu_2\frac{\partial f}{\partial \mu_1} - \mu_1\frac{\partial f}{\partial \mu_2} - \lambda_2\frac{\partial f}{\partial \lambda_1} + \lambda_1\frac{\partial f}{\partial \lambda_2} + 0\frac{\partial f}{\partial \lambda_3}. \tag{6.13}$$

On the other hand, let $J(\xi)$ be a function on M, then

$$\begin{aligned}\{f, J(\xi)\}(\widetilde{z}) =& df(\widetilde{z})^T\widetilde{\Lambda}(\widetilde{z})dJ(\xi)(\widetilde{z})\\ =& \frac{\partial J(\xi)}{\partial \mu_1}\frac{\partial f}{\partial x_1} + \frac{\partial J(\xi)}{\partial \mu_2}\frac{\partial f}{\partial x_2} - \frac{\partial J(\xi)}{\partial x_1}\frac{\partial f}{\partial \mu_1} - \frac{\partial J(\xi)}{\partial x_2}\frac{\partial f}{\partial \mu_2}\\ &+ (\lambda_3\frac{\partial J(\xi)}{\partial \lambda_2} - \lambda_2\frac{\partial J(\xi)}{\partial \lambda_3})\frac{\partial f}{\partial \lambda_1} + (-\lambda_3\frac{\partial J(\xi)}{\partial \lambda_1} + \lambda_1\frac{\partial J(\xi)}{\partial \lambda_3})\frac{\partial f}{\partial \lambda_2}\\ &+ (\lambda_2\frac{\partial J(\xi)}{\partial \lambda_1} - \lambda_1\frac{\partial J(\xi)}{\partial \lambda_2})\frac{\partial f}{\partial \lambda_3}.\end{aligned} \tag{6.14}$$

Comparing (6.13) and (6.14), we have the following PDE

$$\begin{cases} \dfrac{\partial J(\xi)}{\partial \mu_1} = x_2, \\ \dfrac{\partial J(\xi)}{\partial \mu_2} = -x_1, \\ -\dfrac{\partial J(\xi)}{\partial x_1} = \mu_2, \\ -\dfrac{\partial J(\xi)}{\partial x_2} = -\mu_1, \end{cases} \qquad \begin{cases} \lambda_3\dfrac{\partial J(\xi)}{\partial \lambda_2} - \lambda_2\dfrac{\partial J(\xi)}{\partial \lambda_3} = -\lambda_2, \\ -\lambda_3\dfrac{\partial J(\xi)}{\partial \lambda_1} + \lambda_1\dfrac{\partial J(\xi)}{\partial \lambda_3} = \lambda_1, \\ \lambda_2\dfrac{\partial J(\xi)}{\partial \lambda_1} - \lambda_1\dfrac{\partial J(\xi)}{\partial \lambda_2} = 0. \end{cases}$$

One can check that

$$J(\xi)(\widetilde{z}) = -x_1\mu_2 + x_2\mu_1 + \lambda_3$$

is a solution of the above PDE. Therefore, from Noether's theorem, this function is a constant of motion of $X_{H_{12}}$, i.e.,

$$-x_1\mu_2 + x_2\mu_1 + \lambda_3 = C_2 \tag{6.15}$$

for some constant C_2. □

Since the reduced system has G_0-symmetry, by using the standard Poisson reduction procedure again, we can drop the system (6.5) to the quotient space $\bar{P} \triangleq \widetilde{P}/G_0 \simeq T^*(R^2 \times SO(3))/SO(3)/S^1$ with projection $\bar{\pi} : \widetilde{P} \to \bar{P}$. In the following, we will find an induced Hamiltonian $\bar{H}_{12}$, an induced Poisson structure $\bar{\Lambda}$ and a reduced Hamiltonian vector field $X_{\bar{H}_{12}}$ on the manifold $\bar{P}$. First, consider a change of coordinates

$$\begin{gathered}\psi : \widetilde{P} \to \widetilde{P} \\ \widetilde{z} = (x_1, x_2, \mu_1, \mu_2, \lambda_1, \lambda_2, \lambda_3) \mapsto \widetilde{z}' = (r_1, r_2, r_3, \theta_1, \theta_2, \theta_3, \lambda_3)\end{gathered} \tag{6.16a}$$

given by relations

$$\begin{cases} x_1 = r_1\cos(\theta_1), \\ x_2 = r_1\sin(\theta_1), \\ \mu_1 = r_2\cos(\theta_2), \\ \mu_2 = r_2\sin(\theta_2), \end{cases} \qquad \begin{cases} \lambda_1 = r_3\cos(\theta_3), \\ \lambda_2 = -r_3\sin(\theta_3), \\ \lambda_3 = \lambda_3. \end{cases} \tag{6.16b}$$

With these new coordinates, the Hamiltonian H_{12} becomes

$$\begin{aligned} H'_{12}(\widetilde{z}') = & \tfrac{1}{4}(2ar_2r_3\sin(\theta_3-\theta_2) + 2ab\lambda_3r_1r_3\cos(\theta_3-\theta_1) \\ & + 2b\lambda_3r_1r_2\sin(\theta_2-\theta_1) - a^2r_3^2 - r_2^2 - b^2\lambda_3^2r_1^2). \end{aligned} \tag{6.17}$$

And the corresponding dynamics $X_{H'_{12}}$ is given by

$$\begin{cases} \dot{r}_1 = \dfrac{ar_3\sin(\theta_3-\theta_1) - r_2\cos(\theta_2-\theta_1)}{2}, \\ \dot{\theta}_1 = -\dfrac{ar_3\cos(\theta_3-\theta_1) + r_2\sin(\theta_2-\theta_1) - b\lambda_3r_1}{2r_1}, \\ \dot{r}_2 = -\dfrac{ab\lambda_3r_3\cos(\theta_3-\theta_2) - b^2\lambda_3^2r_1\cos(\theta_2-\theta_1)}{2}, \\ \dot{\theta}_2 = -\dfrac{1}{2r_2}(ab\lambda_3r_3\sin(\theta_3-\theta_2) + b^2\lambda_3^2r_1\sin(\theta_2-\theta_1) - b\lambda_3r_2), \\ \dot{r}_3 = \dfrac{a\lambda_3r_2\cos(\theta_3-\theta_2) - ab\lambda_3^2r_1\sin(\theta_3-\theta_1)}{2}, \\ \dot{\theta}_3 = -\dfrac{1}{2r_3}(a\lambda_3r_2\sin(\theta_3-\theta_2) + (ab\lambda_3^2r_1 - abr_1r_3^2)\cos(\theta_3-\theta_1) \\ \qquad - br_1r_2r_3\sin(\theta_2-\theta_1) + (b^2\lambda_3r_1^2 - a^2\lambda_3)r_3), \\ \dot{\lambda}_3 = -\dfrac{ar_2r_3\cos(\theta_3-\theta_2) - ab\lambda_3r_1r_3\sin(\theta_3-\theta_1)}{2}. \end{cases} \tag{6.18}$$

Observing that the Hamiltonian H'_{12} in (6.17) and the right-hand-side of differential equation (6.18) depend on relative values of θ_1, θ_2 and θ_3 only, we can reduce the order of this system as follows. Let

$$\theta_{21} = \theta_2 - \theta_1 \quad \text{and} \quad \theta_{32} = \theta_3 - \theta_2.$$

By using $\bar{z} = (r_1, r_2, r_3, \theta_{21}, \theta_{32}, \lambda_3)$ to parameterize $\bar{P}$, the induced Hamiltonian on $\bar{P}$ is given by

$$\begin{aligned}\bar{H}_{12}(\bar{z}) =& \tfrac{1}{4}(2ab\lambda_3 r_1 r_3 \cos(\theta_{32}+\theta_{21}) + 2ar_2r_3\sin(\theta_{32}) \\ &+ 2b\lambda_3 r_1 r_2 \sin(\theta_{21}) - a^2 r_3^2 - r_2^2 - b^2\lambda_3^2 r_1^2).\end{aligned} \tag{6.19}$$

The corresponding induced dynamics $X_{\bar{H}_{12}}$ on $\bar{P}$ is given by

$$\left\{\begin{aligned}
\dot{r}_1 =& \frac{ar_3\sin(\theta_{32}+\theta_{21}) - r_2\cos(\theta_{21})}{2}, \\
\dot{r}_2 =& -\frac{ab\lambda_3 r_3\cos(\theta_{32}) - b^2\lambda_3^2 r_1\cos(\theta_{21})}{2}, \\
\dot{r}_3 =& -\frac{ab\lambda_3^2 r_1\sin(\theta_{32}+\theta_{21}) - a\lambda_3 r_2\cos(\theta_{32})}{2}, \\
\dot{\theta}_{21} =& \frac{1}{2r_1r_2}(ar_2r_3\cos(\theta_{32}+\theta_{21}) - ab\lambda_3 r_1 r_3\sin(\theta_{32}) \\
&+ (r_2^2 - b^2\lambda_3^2 r_1^2)\sin(\theta_{21})), \\
\dot{\theta}_{32} =& \frac{1}{2r_2r_3}((abr_1r_2r_3^2 - ab\lambda_3^2 r_1 r_2)\cos(\theta_{32}+\theta_{21}) \\
&+ (ab\lambda_3 r_3^2 - a\lambda_3 r_2^2)\sin(\theta_{32}) + (br_1r_2^2 + b^2\lambda_3^2 r_1)r_3\sin(\theta_{21}) \\
&+ ((a^2-b)\lambda_3 - b^2\lambda_3 r_1^2)r_2r_3), \\
\dot{\lambda}_3 =& \frac{ab\lambda_3 r_1 r_3\sin(\theta_{32}+\theta_{21}) - ar_2r_3\cos(\theta_{32})}{2}.
\end{aligned}\right. \tag{6.20}$$

Moreover the first integrals in (5.6) and (6.15) now take the form:

$$r_3^2 + \lambda_3^2 = C_1 \tag{6.21}$$

and

$$r_1 r_2 \sin(\theta_{21}) + \lambda_3 = C_2. \tag{6.22}$$

Therefore, as we claimed before, the final reduced system with the above integrals is a four-dimensional Hamiltonian system.

Remark 6.4 One can further show that the final reduced system (6.20) with (6.21) and (6.22) is also Poisson. Indeed, the equations in (6.20) can be written as

$$\dot{\bar{z}} = \bar{\Lambda}(\bar{z})\nabla\bar{H}_{12}(\bar{z}),$$

where

$$\bar{\Lambda}(\bar{z}) = \begin{pmatrix} 0 & \cos(\theta_{21}) & 0 & -\frac{\sin(\theta_{21})}{r_2} & \frac{\sin(\theta_{21})}{r_2} & 0 \\ -\cos(\theta_{21}) & 0 & 0 & \frac{\sin(\theta_{21})}{r_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\lambda_3}{r_3} \\ \frac{\sin(\theta_{21})}{r_2} & -\frac{\sin(\theta_{21})}{r_1} & 0 & 0 & \frac{\cos(\theta_{21})}{r_1r_2} & 0 \\ -\frac{\sin(\theta_{21})}{r_2} & 0 & 0 & -\frac{\cos(\theta_{21})}{r_1r_2} & 0 & 0 \\ 0 & 0 & -\frac{\lambda_3}{r_3} & 0 & 0 & 0 \end{pmatrix}.$$

For a detailed derivation of the above expression, see [13]. □

7 Final Remarks

The model problem described in this paper is strongly motivated by a troublesome phenomenon of drift observed in the Hubble Space Telescope due to thermoelastically driven vibrations of the solar panels arising from the day-night thermal cycling on-orbit. The point mass oscillators in our problem may be viewed as one-mode truncations of this elasto-mechanical problem.

It should be noted that in Section 4 and 5, the whole analysis does not depend on the contents of the vectors Ω_1 and Ω_2 and the dimension of the shape space is not important either. Therefore, the analysis in that part can be extended to systems with bundle structure $(\mathbb{R}^m \times SO(3), \mathbb{R}^m, SO(3))$, with $m > 1$.

References

[1] P. S. Krishnaprasad, R. Yang, and W. Dayawansa [1991] *Control Problems on Principal Bundles and Nonholonomic Mechanics*, Proceedings of 30th IEEE Conference on Decision and Control, IEEE, New York, pp. 1133–1138.

[2] S. Smale [1970] *Topology and Mechanics I, II*, Invent. Math. **10** 305–331, and **11** 45–64.

[3] J. E. Marsden, R. Montgomery, and T. Ratiu [1990] *Reduction, symmetry, and phases in mechanics*, Memoirs of the AMS.

[4] W. L. Chow [1940] *Ueber Systeme von Linearen Partiellen Differentialgleichungen Erster Ordnung*, Math. Ann. **117**, no. 1, 98–105,.

[5] R. Montgomery [1990] *Isoholonomic Problems and Applications*, Communications in Math. Physics **128**, 565–592.

[6] A. M. Vershik and V. Ya. Gershkovich [1988] *Nonholonomic Problems and the Theory of Distributions*, Acta Applicandae Mathematicae **12**, 181–209.

[7] R. W. Brockett [1982.] *Control Theory and Singular Riemannian Geometry*, New Directions in Applied Mathematics, Springer-Verlag, Berlin,.

[8] L-S. Wang, [1990] *Geometry, Dynamics and Control of Coupled Systems*, Ph.D. Thesis, University of Maryland, College Park.

[9] P. S. Krishnaprasad and R. Yang [1991] *Geometric Phases, Anholonomy, and Optimal Movement*, Proceedings of the IEEE International Conference on Robotics and Automation, IEEE, New York, pp. 2185–2189.

[10] R. Abraham and J.E. Marsden [1978.] *Foundations of Mechanics*, (Second edition), Revised, Enlarged, Reset,, Benjamin/Cummings, Reading,.

[11] J. E. Marsden and A. Weinstein [1974] *Reduction of Symplectic Manifolds with Symmetry*, Reports in Math. Phys., vol. 5, pp. 121–130.

[12] J.E. Marsden and T. Ratiu [1986] *Reduction of Poisson Manifolds*, Letters in Math. Phys., vol. 11, pp. 161–169.

[13] R. Yang [1992.] *Nonholonomic Geometry, Mechanics and Control*, Ph.D. Thesis, University of Maryland, College Park.

Preface

This volume collects the proceedings of a workshop which was held on June 12, 1992 both as a commemoration of the 25th anniversary of the publication of "Foundations of Mechanics" by Ralph Abraham and Jerrold Marsden and as a celebration of Jerry's 50th birthday. The publication of that first edition marked a period of remarkable resurgence in all aspects of mechanics, which has continued, through the publication of the second edition in 1978, deeply nourished by contacts with a variety of areas of mathematics including toplogy, differential geometry, Lie theory, and partial differential equations to name a few. The papers collected in this volume reflect these fruitful ties as well as some others, strengthened over the last two decades, with areas of applied science including control theory.

Jerry Marsden has been involved centrally in many of these developments, through his wide-ranging insights, through his intense collaborations, through his tireless teaching and scholarship, and his creation of new directions of research that have attracted energetic contributions from around the world. Through his personal and intellectual generosity, Jerry Marsden has shaped the subject of geometric mechanics in lasting ways and has been a source of inspiration to us and countless others.

It is with great pleasure that we as editors dedicate this volume to Jerry as a delayed birthday present.