Nonholonomic Mechanics and Control
Internet Supplement

A. M. Bloch

with the collaboration of

J. Baillieul
P. E. Crouch
J. E. Marsden

with scientific input from

P. S. Krishnaprasad
R. M. Murray
D. Zenkov

This Version: May 19, 2007

Preface

The purpose of this supplement is to give some additional topics not covered in the main text as well as to provide some additional details and sometimes, alternative approaches. For example, §2.9 provides the proof of the Cartan structure equations which were just stated in the text and the supplement to §5.3 gives an alternative approach to the geometric foundations of nonholonomic systems.

The supplement is organized using the same structure as the book itself, namely material most relevant to a section in the book is given the same section number in this supplement.

This supplement will be updated from time to time, so check back for updates and please do let us know of any comments or corrections.

Contents

2

Mathematical Preliminaries

2.1 Vector Fields, Flows, and Differential Equations

Existence of Solutions for a Particle in a Potential Field. This supplement gives some conditions based on energy estimates, which guarantee the existence of long time solutions to the dynamics of a particle in a potential field.[1]

Consider a particle of mass m moving in $\mathbb{R}^n$ in a potential field $V : \mathbb{R}^n \to \mathbb{R}$. The equations are, from Newton's second law,

$$\ddot{\mathbf{q}}(t) = -(1/m)\nabla V(\mathbf{q}(t)).$$

We shall prove that *if there are constants* $a, b \in \mathbb{R}, b \geq 0$ *such that*

$$\frac{1}{m} V(\mathbf{q}) \geq a - b\|\mathbf{q}\|^2,$$

then every solution exists for all time. To show this, rewrite the second order equations as a first order system $\dot{\mathbf{q}} = (1/m)\mathbf{p}, \dot{\mathbf{p}} = -\nabla V(\mathbf{q})$ and note that the energy

$$E(\mathbf{q}, \mathbf{p}) = \frac{1}{2m}\|\mathbf{p}\|^2 + V(\mathbf{q})$$

is a first integral. Thus, for any solution $(\mathbf{q}(t), \mathbf{p}(t))$ we have

$$\beta = E(\mathbf{q}(t), \mathbf{p}(t)) = E(\mathbf{q}(0), \mathbf{p}(0)) \geq V(\mathbf{q}(0)).$$

[1] This exposition is adapted from Marsden and Weinstein [1970].

We can assume $\beta > V(\mathbf{q}(0))$, *i.e.,* $\mathbf{p}(0) \neq 0$, for if $\mathbf{p}(t) \equiv 0$, then the conclusion is trivially satisifed; thus there exists a t_0 for which $\mathbf{p}(t_0) \neq 0$ and by time translation we can assume that $t_0 = 0$. Thus we have

$$\begin{aligned}\|\mathbf{q}(t)\| \leq \|\mathbf{q}(t) - \mathbf{q}(0)\| + \|\mathbf{q}(0)\| &\leq \|\mathbf{q}(0)\| + \int_0^t \|\dot{\mathbf{q}}(s)\|\, ds \\ &= \|\mathbf{q}(0)\| + \int_0^t \sqrt{2\left[\beta - \frac{1}{m}V(\mathbf{q}(s))\right]}\, ds \\ &\leq \|\mathbf{q}(0)\| + \int_0^t \sqrt{2(\beta - a + b\|\mathbf{q}(s)\|^2)}\, ds\end{aligned}$$

or in differential form

$$\frac{d}{dt}\|\mathbf{q}(t)\| \leq \sqrt{2(\beta - a + b\|\mathbf{q}(t)\|^2)}$$

whence

$$t \geq \int_{\|\mathrm{q}(0)\|}^{\|\mathrm{q}(t)\|} \frac{du}{\sqrt{2(\beta - a + bu^2)}}\,. \tag{2.1.3}$$

Now let $r(t)$ be the solution of the differential equation

$$\frac{d^2 r(t)}{dt^2} = -\frac{d}{dr}(a - br^2)(t) = 2br(t),$$

which, as a second order equation with constant coefficients, has solutions for all time for any initial conditions. Choose

$$r(0) = \|\mathbf{q}(0)\|, [\dot{r}(0)]^2 = 2(\beta - a + b\|\mathbf{q}(0)\|^2)$$

and let $r(t)$ be the corresponding solution. Since

$$\frac{d}{dt}\left(\frac{1}{2}\dot{r}(t)^2 + a - br(t)^2\right) = 0,$$

it follows that $(1/2)\dot{r}(t)^2 + a - br(t)2 = (1/2)\dot{r}(0)^2 + a - br(0)^2 = \beta$, *i.e.,*

$$\frac{dr(t)}{dt} = \sqrt{2(\beta - a + br(t)^2)}$$

whence

$$t = \int_{\|\mathrm{q}(0)\|}^{r(t)} \frac{du}{\sqrt{2(\beta - a + bu^2}}\,. \tag{2.1.4}$$

Comparing these two expressions and taking into account that the integrand is > 0, it follows that for any finite time interval for which $\mathbf{q}(t)$ is defined, we have $\|\mathbf{q}(t)\| \leq r(t)$, *i.e.,* $\mathbf{q}(t)$ remains in a compact set for finite t-intervals. But then $\dot{\mathbf{q}}(t)$ also lies in a compact set since

$$\|\dot{\mathbf{q}}(t)\| \leq 2(\beta - a + b\|\mathbf{q}(s)\|^2).$$

Thus by Proposition 2.1.11 (a criterion for completeness) of the main book, the solution curve $(\mathbf{q}(t), \mathbf{p}(t))$ is defined for any $t \geq 0$. However, since $(\mathbf{q}(-t), \mathbf{p}(-t))$ is the value at t of the integral curve with initial conditions $(-\mathbf{q}(0), -\mathbf{p}(0))$, it follows that the solution also exists for all $t \leq 0$.

The following counterexample shows that the condition $V(\mathbf{q}) \geq a - b\|\mathbf{q}\|^2$ cannot be relaxed much further. Take $n = 1$ and $V(q) = -\varepsilon^2 q^{2+(4/\varepsilon)}/8, \varepsilon > 0$. Then the equation

$$\ddot{q} = \varepsilon(\varepsilon + 2)q^{1+(4/\varepsilon)}/4$$

has the solution $q(t) = 1/(t-1)^{\varepsilon/2}$, which cannot be extended beyond $t = 1$.

2.9 Fiber Bundles and connections

In this supplement we provide a number of additional topics and details on principal connections.

The Cartan Structure Equations. We begin with the details of the proof of the Cartan structure equations, which were just stated in the text. We use the notation and set up given in the text.

2.9.13 Theorem (Cartan Structure Equations). *For any vector fields u, v on Q we have*

$$\mathcal{B}(u, v) = \mathbf{d}\mathcal{A}(u, v) - [\mathcal{A}(u), \mathcal{A}(v)]$$

where the bracket on the right hand side is the Lie bracket in $\mathfrak{g}$. We write this equation for short as

$$\mathcal{B} = \mathbf{d}\mathcal{A} - [\mathcal{A}, \mathcal{A}].$$

To prove this theorem we prepare a lemma.

Lemma. *We have the identity* $\mathbf{d}\mathcal{A}(\mathrm{hor}(u), \mathrm{ver}(v)) = 0$ *for any two vector fields u, v on Q .*

Proof. Since this identity depends only on the point values of u and v, we can assume that $\mathrm{ver}(v) = \xi_Q$ identically. Then, as in the preceding proposition, we have

$$\begin{aligned}
\mathbf{d}\mathcal{A}(\mathrm{hor}(u), \mathrm{ver}(v)) &= (\mathrm{hor}(u))[\mathcal{A}(\xi_Q)] - \xi_Q[\mathcal{A}(\mathrm{hor}(u))] - \mathcal{A}([\mathrm{hor}(u), \xi_Q]) \\
&= \mathrm{hor}(u)[\xi] - \xi_Q[0] + \mathcal{A}[\xi_Q, \mathrm{hor}(u)] \\
&= \mathcal{A}[\xi_Q, \mathrm{hor}(u)]
\end{aligned}$$

since ξ is constant. However, the flow of ξ_Q is $\Phi_{\exp(t\xi)}$ and the map hor is equivariant and so

$$\begin{aligned}[\xi_Q, \operatorname{hor}(u)] &= \left.\frac{d}{dt}\right|_{t=0} \Phi^*_{\exp(t\xi)} \operatorname{hor}(u) \\ &= \operatorname{hor} \left.\frac{d}{dt}\right|_{t=0} \Phi^*_{\exp(t\xi)}(u) \\ &= \operatorname{hor}[\xi_Q, u]\end{aligned}$$

Thus, $[\xi_Q, \operatorname{hor}(u)]$ is horizontal and so it is annihilated by $\mathcal{A}$ and so the lemma follows. ■

Proof of the Cartan structure equations. Use of the lemma and writing $u = \operatorname{hor}(u) + \operatorname{ver}(u)$ and similarly for v, shows that

$$\mathbf{d}\mathcal{A}(u, \operatorname{ver}(v)) = \mathbf{d}\mathcal{A}(\operatorname{ver}(u), \operatorname{ver}(v))$$

and so we get

$$\mathcal{B}(u, v) = \mathbf{d}\mathcal{A}(u, v) - \mathbf{d}\mathcal{A}(\operatorname{ver}(u), \operatorname{ver}(v)).$$

Again, the second term on the right hand side of this equation depends only on the point values of u and v and so we can assume that $\operatorname{ver}(u) = \xi_Q$ and that $\operatorname{ver}(v) = \eta_Q$ for $\xi \in \mathfrak{g}$ and $\eta \in \mathfrak{g}$. Then

$$\begin{aligned}\mathbf{d}\mathcal{A}(\xi_Q, \eta_Q) &= \xi_Q[\mathcal{A}(\eta_Q)] - \eta_Q[\mathcal{A}(\xi_Q)] - \mathcal{A}([\xi_Q, \eta_Q]) \\ &= \mathcal{A}([\xi, \eta]_Q) = [\xi, \eta] \\ &= [\mathcal{A}(u), \mathcal{A}(v)].\end{aligned}$$ ■

The following Corollary shows how the Cartan Structure Equations yield a fundamental equivariance property of the curvature.

2.9.14 Corollary. *For all $g \in G$ we have $\Phi_g^*\mathcal{B} = \operatorname{Ad}_g \circ \mathcal{B}$. If the G-action on Q is on the right, equivariance means $\Phi_g^*\mathcal{B} = \operatorname{Ad}_{g^{-1}} \circ \mathcal{B}$.*

Proof. From the definition of $[\mathcal{A}, \mathcal{A}]$, we get for any $u_q, v_q \in T_qQ$ and $g \in G$

$$\begin{aligned}\left(\Phi_g^*[\mathcal{A}, \mathcal{A}]\right)(q)(u_q, v_q) &= [\mathcal{A}, \mathcal{A}](g \cdot q)\left(T_q\Phi_g(u_q), T_q\Phi_g(v_q)\right) \\ &= [\mathcal{A}(q)\left(T_q\Phi_g(u_q)\right), \mathcal{A}(q)\left(T_q\Phi_g(v_q)\right)] \\ &= [\left(\Phi_g^*\mathcal{A}\right)(q)(u_q), \left(\Phi_g^*\mathcal{A}\right)(q)(v_q)] \\ &= [\Phi_g^*\mathcal{A}, \Phi_g^*\mathcal{A}](q)(u_q, v_q),\end{aligned}$$

that is, $\Phi_g^*[\mathcal{A},\mathcal{A}] = [\Phi_g^*\mathcal{A},\Phi_g^*\mathcal{A}]$. Thus equivariance of the connection gives

$$\Phi_g^*[\mathcal{A},\mathcal{A}] = [\Phi_g^*\mathcal{A},\Phi_g^*\mathcal{A}] = [\mathrm{Ad}_g \circ\mathcal{A}, \mathrm{Ad}_g \circ\mathcal{A}] = \mathrm{Ad}_g \circ[\mathcal{A},\mathcal{A}]$$

and hence

$$\begin{aligned}\Phi_g^*\mathcal{B} &= \Phi_g^*\left(\mathbf{d}\mathcal{A} - [\mathcal{A},\mathcal{A}]\right) = \mathbf{d}\left(\Phi_g^*\mathcal{A}\right) - \mathrm{Ad}_g \circ[\mathcal{A},\mathcal{A}]\\ &= \mathbf{d}\left(\mathrm{Ad}_g \circ\mathcal{A}\right) - \mathrm{Ad}_g \circ[\mathcal{A},\mathcal{A}] = \mathrm{Ad}_g \circ\left(\mathbf{d}\mathcal{A} - [\mathcal{A},\mathcal{A}]\right) = \mathrm{Ad}_g \circ\mathcal{B}\end{aligned}$$

as required. The case of right actions is proved in a similar way. ■

Curvature as a Two-Form on the Base. We now show how the curvature two-form drops to a two-form on the base with values in a bundle called the adjoint bundle.

The associated bundle to the given left principal bundle $\pi_{Q,G} : Q \to Q/G$ via the adjoint action is called the ***adjoint bundle*** and is defined as follows. Consider the free proper action

$$(g,(q,\xi)) \in G \times (Q \times \mathfrak{g}) \mapsto (g \cdot q, \mathrm{Ad}_g \xi) \in Q \times \mathfrak{g}$$

and form the quotient

$$\tilde{\mathfrak{g}} := Q \times_G \mathfrak{g} := (Q \times \mathfrak{g})/G,$$

which is verified to be a vector bundle $\pi_{\tilde{\mathfrak{g}}} : \tilde{\mathfrak{g}} \to Q/G$, where $\pi_{\tilde{\mathfrak{g}}}(g,\xi) := \pi_{Q,G}(q)$. This vector bundle has an additional structure: it is a ***Lie algebra bundle***; that is, a vector bundle whose fibers are Lie algebras. In this case the bracket is defined pointwise:

$$[\pi_{\tilde{\mathfrak{g}}}(g,\xi), \pi_{\tilde{\mathfrak{g}}}(g,\eta)] := \pi_{\tilde{\mathfrak{g}}}(g,[\xi,\eta])$$

for all $g \in G$ and $\xi,\eta \in \mathfrak{g}$. It is easy to check that this defines a Lie bracket on every fiber and that this operation is smooth as a function of $\pi_{Q,G}(q)$.

The curvature two-form $\mathcal{B} \in \Omega^2(Q;\mathfrak{g})$ (the vector space of $\mathfrak{g}$-valued two-forms on Q) naturally induces a two-form $\overline{\mathcal{B}}$ on the base Q/G with values in $\tilde{\mathfrak{g}}$ by

$$\overline{\mathcal{B}}(\pi_{Q,G}(q))\left(T_q\pi_{Q,G}(u), T_q\pi_{Q,G}(v)\right) := \pi_{\tilde{\mathfrak{g}}}\left(q, \mathcal{B}(u,v)\right) \tag{2.9.1}$$

for all $q \in Q$ and $u,v \in T_qQ$.

We need to check that $\overline{\mathcal{B}}$ is well defined. If $q' = g \cdot q$ and $u',v' \in T_{q'}Q$ are such that

$$T_{q'}\pi_{Q,G}(u') = T_{q'}\pi_{Q,G}\left(T_q\Phi_g(u)\right) = T_q\pi_{Q,G}(u)$$

and

$$T_{q'}\pi_{Q,G}(v') = T_{q'}\pi_{Q,G}\left(T_q\Phi_g(v)\right) = T_q\pi_{Q,G}(v),$$

then $u' - T_q\Phi_g(u), v' - T_q\Phi_g(v) \in V_{q'}$ and hence by Corollary 2.9.14 we get

$$\begin{aligned}
&\pi_{\tilde{\mathfrak{g}}}\left(q',\, \mathcal{B}(u',v')\right) \\
&\qquad = \pi_{\tilde{\mathfrak{g}}}\left(g\cdot q,\, \mathcal{B}(T_q\Phi_g(u) + (u' - T_q\Phi_g(u)), T_q\Phi_g(v) + (v' - T_q\Phi_g(v))\right) \\
&\qquad = \pi_{\tilde{\mathfrak{g}}}\left(g\cdot q,\, \mathcal{B}(T_q\Phi_g(u), T_q\Phi_g(v))\right) \\
&\qquad = \pi_{\tilde{\mathfrak{g}}}\left(g\cdot q,\, (\Phi_g^*\mathcal{B})(u,v)\right) \\
&\qquad = \pi_{\tilde{\mathfrak{g}}}\left(g\cdot q,\, \mathrm{Ad}_g\left(\mathcal{B}(u,v)\right)\right) \\
&\qquad = \pi_{\tilde{\mathfrak{g}}}\left(q,\, \mathcal{B}(u,v)\right)
\end{aligned}$$

which shows that the right hand side of (2.9.1) is independent of the choices made to define $\overline{\mathcal{B}}$.

Since (2.9.1) can be equivalently written as $\pi_{Q,G}^*\overline{\mathcal{B}} = \pi_{\tilde{\mathfrak{g}}} \circ (\mathrm{id}_Q \times \mathcal{B})$ and $\pi_{Q,G}$ is a surjective submersion, it follows that $\overline{\mathcal{B}}$ is indeed a smooth two-form on Q/G with values in $\tilde{\mathfrak{g}}$.

Associated One-Forms. Since $\mathcal{A}$ is a Lie algebra valued 1-form, for each $q \in Q$, we get a linear map $\mathcal{A}(q) : T_qQ \to \mathfrak{g}$ and so we can form its dual $\mathcal{A}(q)^* : \mathfrak{g}^* \to T_q^*Q$. Evaluating this on $\mu \in \mathfrak{g}^*$ produces an ordinary 1-form:

$$\alpha_\mu(q) = \mathcal{A}(q)^*(\mu). \tag{2.9.2}$$

This 1-form satisfies two important properties given in the next Proposition. (Here $\mathbf{J}$ is the cotangent momentum map that is discussed in §3.7 of the main text).

2.9.15 Proposition. *For any connection $\mathcal{A}$ and $\mu \in \mathfrak{g}^*$, the corresponding 1-form α_μ defined by (2.9.2) takes values in $\mathbf{J}^{-1}(\mu)$ and satisfies the following G-equivariance property:*

$$\Phi_g^*\alpha_\mu = \alpha_{\mathrm{Ad}_g^*\mu}.$$

Proof. First of all, notice that from the definition of α_μ and then using first property of a connection,

$$\begin{aligned}
\langle \mathbf{J}(\alpha_\mu(q)), \xi\rangle &= \langle \alpha_\mu(q), \xi_Q(q)\rangle \\
&= \langle \mathcal{A}(q)^*(\mu), \xi_Q(q)\rangle \\
&= \langle \mu, \mathcal{A}(q)(\xi_Q(q))\rangle \\
&= \langle \mu, \xi\rangle\,.
\end{aligned}$$

Since $\xi \in \mathfrak{g}$ is arbitrary, we conclude that $\mathbf{J}(\alpha_\mu(q)) = \mu$ and therefore, indeed, α_μ takes values in $\mathbf{J}^{-1}(\mu)$.

To establish invariance of the form α_μ, we compute in the following way. Let $v \in T_qQ$ and $g \in G$, and first use the definition of α_μ and the definition of the adjoint to get

$$\begin{aligned}(\Phi_g^*\alpha_\mu)(v) &= \alpha_\mu(g\cdot q)(T_q\Phi_g(v))\\ &= \langle \mathcal{A}(g\cdot q)^*(\mu), T_q\Phi_g(v)\rangle\\ &= \langle \mu, \mathcal{A}(g\cdot q)(T_q\Phi_g(v))\rangle\,.\end{aligned}$$

Next, make use of equivariance of $\mathcal{A}$ and convert the preceding expression back to one involving α_μ to get:

$$\begin{aligned}(\Phi_g^*\alpha_\mu)(v) &= \langle \mu, \mathrm{Ad}_g(\mathcal{A}(q)(v))\rangle\\ &= \left\langle \mathrm{Ad}_g^*\mu, \mathcal{A}(q)(v)\right\rangle\\ &= \langle \mathcal{A}(q)^*\mathrm{Ad}_g^*\mu, v\rangle\\ &= \alpha_{\mathrm{Ad}_g^*\mu}(q)(v)\end{aligned}$$

so that we get the required equivariance property. ■

Notice in particular, if the group is Abelian or if μ is G-invariant, (for example, if $\mu = 0$), then α_μ is an *invariant* 1*-form.*

Associated Two-Forms. Since $\mathcal{B}$ is a $\mathfrak{g}$-valued two-form, in analogy with (2.9.2), for every $\mu \in \mathfrak{g}^*$ we can define the ***μ-component*** of $\mathcal{B}$, an ordinary two-form $\mathcal{B}_\mu \in \Omega^2(Q)$ on Q, by

$$\mathcal{B}_\mu(q)(u_q, v_q) := \langle \mu, \mathcal{B}(q)(u_q, v_q)\rangle \tag{2.9.3}$$

for all $q \in Q$ and $u_q, v_q \in T_qQ$.

The adjoint bundle valued curvature two-form $\overline{\mathcal{B}}$ induces an ordinary two-form on the base Q/G. To obtain it, we consider the dual $\tilde{\mathfrak{g}}^*$ of the adjoint bundle. This is a vector bundle over Q/G which is the associated bundle relative to the coadjoint action of the structure group G of the principal (left) bundle $\pi_{Q,G} : Q \to Q/G$ on $\mathfrak{g}^*$. This vector bundle has additional structure: each of its fibers is a Lie-Poisson space and the associated Poisson tensors on each fiber depend smoothly on the base, that is, $\pi_{\tilde{\mathfrak{g}}^*} : \tilde{\mathfrak{g}}^* \to Q/G$ is a ***Lie-Poisson bundle*** over Q/G.

Given $\mu \in \mathfrak{g}^*$, define the *ordinary two-form* $\overline{\mathcal{B}}_\mu$ *on* Q/G by

$$\begin{aligned}&\overline{\mathcal{B}}_\mu\left(\pi_{Q,G}(q)\right)\left(T_q\pi_{Q,G}(u_q), T_q\pi_{Q,G}(v_q)\right)\\ &\quad := \left\langle \pi_{\tilde{\mathfrak{g}}^*}(q,\mu), \overline{\mathcal{B}}(\pi_{Q,G}(q))\left(T_q\pi_{Q,G}(u_q), T_q\pi_{Q,G}(v_q)\right)\right\rangle\\ &\quad = \langle \mu, \mathcal{B}(q)(u_q, v_q)\rangle = \mathcal{B}_\mu(q)(u_q, v_q),\end{aligned} \tag{2.9.4}$$

where $q \in Q$, $u_q, v_q \in T_qQ$, and in the second equality $\langle\,,\rangle : \tilde{\mathfrak{g}}^* \times \tilde{\mathfrak{g}} \to \mathbb{R}$ is the duality pairing between the coadjoint and adjoint bundles. Since $\overline{\mathcal{B}}$ is well defined and smooth, so is $\overline{\mathcal{B}}_\mu$.

2.9.16 Proposition. *Let $\mathcal{A} \in \Omega^1(Q;\mathfrak{g})$ be a connection one-form on the (left) principal bundle $\pi_{Q,G} : Q \to Q/G$ and $\mathcal{B} \in \Omega^2(Q;\mathfrak{g})$ its curvature* *two-form on Q. If $\mu \in \mathfrak{g}^*$, the corresponding two-forms $\mathcal{B}_\mu \in \Omega^2(Q)$ and $\overline{\mathcal{B}}_\mu \in \Omega^2(Q/G)$ defined by (2.9.3) and (2.9.4), respectively, are related by $\pi^*_{Q,G}\overline{\mathcal{B}}_\mu = \mathcal{B}_\mu$. In addition, $\mathcal{B}_\mu$ satisfies the following G-equivariance property:*

$$\Phi_g^* \mathcal{B}_\mu = \mathcal{B}_{\mathrm{Ad}_g^* \mu}.$$

Thus, if $G = G_\mu$ then $\mathbf{d}\alpha_\mu = \mathcal{B}_\mu = \pi^*_{Q,G}\overline{\mathcal{B}}_\mu$, where $\alpha_\mu(q) = \mathcal{A}(q)^*(\mu)$.

Proof. The identity $\pi^*_{Q,G}\overline{\mathcal{B}}_\mu = \mathcal{B}_\mu$ is a restatement of (2.9.4). To prove the equivariance of $\mathcal{B}_\mu$, note that for $g \in G$, Corollary 2.9.14 yields

$$\Phi_g^* \mathcal{B}_\mu = \left\langle \mu, \Phi_g^* \mathcal{B} \right\rangle = \langle \mu, \mathrm{Ad}_g \circ \mathcal{B} \rangle = \left\langle \mathrm{Ad}_g^* \mu, \mathcal{B} \right\rangle = \mathcal{B}_{\mathrm{Ad}_g^* \mu}$$

as required.

The last relation is a consequence of Proposition 2.9.15. Indeed, if $G = G_\mu$ then $\Phi_g^* \alpha_\mu = \alpha_\mu$ for any $g \in G$ so taking the derivative of this relation relative to g at the identity yields $\boldsymbol{\pounds}_{\xi_Q} \alpha_\mu = 0$. However, we also know that α_μ takes values in $\mathbf{J}^{-1}(\mu)$ and hence

$$(\mathbf{i}_{\xi_Q} \alpha_\mu)(q) = \langle \alpha_\mu(q), \xi_Q(q) \rangle = \langle \mathbf{J}(\alpha_\mu(q)), \xi \rangle = \langle \mu, \xi \rangle \,,$$

that is, $\mathbf{i}_{\xi_Q} \alpha_\mu$ is a constant function on Q. Therefore,

$$\mathbf{i}_{\xi_Q} \mathbf{d}\alpha_\mu = \boldsymbol{\pounds}_{\xi_Q} \alpha_\mu - \mathbf{d}\mathbf{i}_{\xi_Q} \alpha_\mu = 0.$$

Now let $q \in Q$, $u_q, v_q \in T_qQ$. Then $u_q - [\langle \mathcal{A}(q), u_q \rangle]_Q(q)$ is the horizontal component of u_q and similarly for v_q. Therefore,

$$\begin{aligned} \mathcal{B}_\mu(q)(u_q, v_q) &= \langle \mu,\, \mathbf{d}\mathcal{A}(q)\,(u_q - [\langle \mathcal{A}(q), u_q \rangle]_Q(q),\, v_q - [\langle \mathcal{A}(q), u_q \rangle]_Q(q)) \rangle \\ &= \mathbf{d}\mathcal{A}(q)\,(u_q, v_q)\,, \end{aligned}$$

since each of the remaining three terms in the expansion is of the form $\mathbf{i}_{\xi_Q} \mathbf{d}\alpha_\mu = 0$, as was shown above. ■

3

Basic Concepts in geometric Mechanics

3.7 Momentum Maps.

Singularities and Symmetry. The following is a basic link between symmetries and singularities.

3.7.8 Proposition. *Let (P, Ω) be a symplectic manifold, let G act on P by Poisson mappings, and let $\mathbf{J} : P \to \mathfrak{g}^*$ be a momentum map for this action ($\mathbf{J}$ need not be equivariant). Let G_z denote the symmetry group of $z \in P$ defined by $G_z = \{g \in G \mid gz = z\}$ and let $\mathfrak{g}_z$ be its Lie algebra, so $\mathfrak{g}_z = \{\zeta \in \mathfrak{g} \mid \zeta_P(z) = 0\}$. Then z is a regular value of $\mathbf{J}$ if and only if $\mathfrak{g}_z$ is trivial; i.e., $\mathfrak{g}_z = \{0\}$, or G_z is discrete.*

Proof. The point z is regular when the range of the linear map $\mathbf{DJ}(z)$ is all of $\mathfrak{g}^*$. However, $\zeta \in \mathfrak{g}$ is orthogonal to the range (in the sense of the $\mathfrak{g}, \mathfrak{g}^*$ pairing) if and only if for all $v \in T_zP$,

$$\langle \zeta, \mathbf{DJ}(z) \cdot v \rangle = 0$$

i.e.,

$$\mathbf{d}\langle \mathbf{J}, \zeta \rangle(z) \cdot v = 0,$$

which is equivalent to

$$\Omega(X_{\langle \mathbf{J}, \zeta \rangle}(z), v) = 0,$$

that is,

$$\Omega(\zeta_P(z), v) = 0.$$

As Ω is nondegenerate, ζ is orthogonal to the range iff $\zeta_P(z) = 0$. ∎

The above proposition is due to Smale [1970]. It is the starting point of a large literature on singularities in the momentum map and singular reduction. Arms, Marsden, and Moncrief [1981] show, under some reasonable hypotheses, that the level sets $\mathbf{J}^{-1}(0)$ have *quadratic* singularities. In the finite dimensional case, this result can be deduced from the equivariant Darboux theorem.[1]

The ***convexity theorem*** states that the image of the momentum map of a torus action is a convex polyhedron in $\mathfrak{g}^*$; the boundary of the polyhedron is the image of the singular (symmetric) points in P; the more symmetric the point, the more singular the boundary point. These results are due to Atiyah [1982] and Guillemin and Sternberg [1984] based on earlier convexity results of Kostant and the Shur-Horn theorem on eigenvalues of symmetric matrices. The literature on these topics and its relation to other areas of mathematics is vast. See, for example, Goldman and Millson [1990], Bloch, Flaschka, and Ratiu [1990], Bloch, Brockett, and Ratiu [1992], Sjamaar and Lerman [1991] and Lu and Ratiu [1991].

3.8 Symplectic and Poisson Reduction.

The text developed some of the basics of reduction theory. An important ingredient that ties this theory to the theory of connections in §2.9 is that of the ***mechanical connection*** which we develop in this supplement.

The Mechanical Connection. As an example of defining a connection by the specification of a horizontal space, suppose that the configuration manifold Q is a Riemannian manifold. Of course, the Riemannian structure will often be that defined by the kinetic energy of a given mechanical system.

Thus, assume that Q is a Riemannian manifold, with metric denoted $\langle\!\langle \, , \rangle\!\rangle$ and that G acts freely and properly on Q by isometries, so $\pi_{Q,G} : Q \to Q/G$ is a principal G-bundle.

In this context we may *define the horizontal space at a point simply to be the metric orthogonal to the vertical space*. This therefore defines a connection called the ***mechanical connection***.

Recall from the historical survey in the introduction that this connection was first introduced by Kummer [1981] following motivation from Smale [1970] and Abraham and Marsden [1978]. See also Guichardet [1984], who applied these ideas in an interesting way to molecular dynamics. The num-

[1] In the infinite dimensional case, things are more subtle. In fact, the infinite dimensional results were motivated by, and apply to, the singularities in the solution space of relativistic field theories such as gravity and the Yang-Mills equations (see Fischer, Marsden, and Moncrief [1980], Arms, Marsden, and Moncrief [1981, 1982] and Arms [1981]).

ber of references since then making use of the mechanical connection is too large to survey here.

In Proposition 3.8.5 we develop an explicit formula for the associated Lie algebra valued 1-form in terms of an inertia tensor and the momentum map. As a prelude to this formula, we show the following basic link with mechanics. In this context we write the momentum map on TQ simply as $\mathbf{J} : TQ \to \mathfrak{g}^*$.

3.8.4 Proposition. *The horizontal space of the mechanical connection at a point $q \in Q$ consists of the set of vectors $v_q \in T_qQ$ such that $\mathbf{J}(v_q) = 0$.*

Proof. This follows directly from the formula for the momentum map for a Lagrangian that is given by the kinetic energy of a given Riemannian metric, namely (see equation (3.7.4) of the text),

$$\langle \mathbf{J}(v_q), \xi \rangle = \langle\!\langle v_q, \xi_Q(q) \rangle\!\rangle$$

and the fact that the vertical space at $q \in Q$ is spanned by the set of infinitesimal generators $\xi_Q(q)$. ■

For each $q \in Q$, define the ***locked inertia tensor*** $\mathbb{I}(q)$ to be the linear map $\mathbb{I}(q) : \mathfrak{g} \to \mathfrak{g}^*$ defined by

$$\langle \mathbb{I}(q)\eta, \zeta \rangle = \langle\!\langle \eta_Q(q), \zeta_Q(q) \rangle\!\rangle \tag{3.8.6}$$

for any $\eta, \zeta \in \mathfrak{g}$. Since the action is free, $\mathbb{I}(q)$ is nondegenerate, so (3.8.6) defines an inner product. The terminology "locked inertia tensor" comes from the fact that for coupled rigid or elastic systems, $\mathbb{I}(q)$ is the classical moment of inertia tensor of the rigid body obtained by locking all the joints of the system. In coordinates,

$$I_{ab} = g_{ij} K^i_a K^j_b, \tag{3.8.7}$$

where $[\xi_Q(q)]^i = K^i_a(q)\xi^a$ define the ***action functions*** K^i_a.

Define the map $\mathcal{A} : TQ \to \mathfrak{g}$ which assigns to each $v_q \in T_qQ$ the corresponding *angular velocity of the locked system*:

$$\mathcal{A}(q)(v_q) = \mathbb{I}(q)^{-1}(\mathbf{J}(\mathbb{F}L(v_q))), \tag{3.8.8}$$

where L is the kinetic energy Lagrangian. In coordinates,

$$\mathcal{A}^a = I^{ab} g_{ij} K^i_b v^j \tag{3.8.9}$$

since $J_a(q, p) = p_i K^i_a(q)$.

We defined the mechanical connection by declaring its horizontal space to be the metric orthogonal to the vertical space. The next proposition shows that $\mathcal{A}$ is the associated connection one-form.

3.8.5 Proposition. *The $\mathfrak{g}$-valued one-form defined by (3.8.8) is the mechanical connection on the principal G-bundle $\pi_{Q,G} : Q \to Q/G$.*

Proof. First notice that $\mathcal{A}$ is G-equivariant and satisfies $\mathcal{A}(\xi_Q(q)) = \xi$, both of which are readily verified. In checking equivariance, one uses invariance of the metric and hence equivariance of $\mathbb{F}L : TQ \to T^*Q$, where L is the kinetic energy of the metric, equivariance of $\mathbf{J} : T^*Q \to \mathfrak{g}^*$, and equivariance of $\mathbb{I}$ in the sense of a map $\mathbb{I} : Q \times \mathfrak{g} \to \mathfrak{g}^*$; that is,

$$\mathbb{I}(g \cdot q)(\mathrm{Ad}_g \xi) = \mathrm{Ad}^*_{g^{-1}} \mathbb{I}(q)(\xi).$$

Thus, $\mathcal{A}$ is a connection.

The horizontal space of $\mathcal{A}$ is given by

$$H_q = \{v_q \in T_qQ \mid \mathbf{J}(\mathbb{F}L(v_q)) = 0\}. \tag{3.8.10}$$

Thus, by Proposition 3.8.4 and the fact that any two connections with the same horizontal spaces are equal, we get the result. ■

Given a general connection $\mathcal{A}$ and an element $\mu \in \mathfrak{g}^*$, we can define the μ-component of $\mathcal{A}$ to be the ordinary one-form α_μ given by

$$\alpha_\mu(q) = \mathcal{A}(q)^*\mu \in T^*_qQ; \quad \text{i.e.,} \quad \langle \alpha_\mu(q), v_q \rangle = \langle \mu, \mathcal{A}(q)(v_q) \rangle$$

for all $v_q \in T_qQ$. Note that α_μ is a G_μ-invariant one-form. It takes values in $\mathbf{J}^{-1}(\mu)$ since for any $\xi \in \mathfrak{g}$, we have

$$\langle \mathbf{J}(\alpha_\mu(q)), \xi \rangle = \langle \alpha_\mu(q), \xi_Q \rangle = \langle \mu, \mathcal{A}(q)(\xi_Q(q)) \rangle = \langle \mu, \xi \rangle .$$

In the case of the mechanical connection, Smale [1970] constructed α_μ by a minimization process. Let $\alpha_q^\sharp \in T_qQ$ be the tangent vector that corresponds to $\alpha_q \in T^*_qQ$ via the metric $\langle\!\langle\, , \rangle\!\rangle$ on Q.

3.8.6 Proposition. *The 1-form $\alpha_\mu(q) = \mathcal{A}(q)^*\mu \in T^*_qQ$ associated with the mechanical connection $\mathcal{A}$ given by (3.8.8) is characterized by*

$$K(\alpha_\mu(q)) = \inf\{K(\beta_q) \mid \beta_q \in \mathbf{J}^{-1}(\mu) \cap T^*_qQ\}, \tag{3.8.11}$$

where $K(\beta_q) = \frac{1}{2}\|\beta_q^\sharp\|^2$ is the kinetic energy function on T^*Q. See Figure 3.8.3.

The proof is a direct verification. We do not give here it since this proposition will not be used later in this book. The original approach of Smale [1970] was to take (3.8.11) as the definition of α_μ. To prove from here that α_μ is a smooth one-form is a nontrivial fact; see the proof in Smale [1970] or of Proposition 4.4.5 in Abraham and Marsden [1978]. Thus, one of the merits of the previous proposition is to show easily that this variational definition of α_μ does indeed yield a smooth one-form on Q with

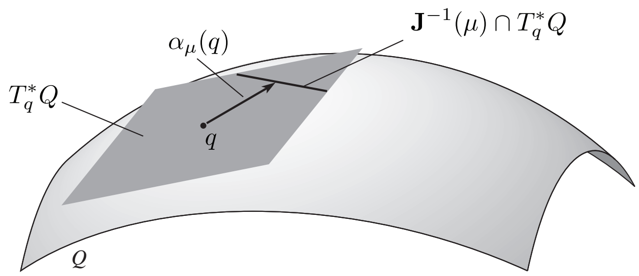

FIGURE 3.8.3. The extremal characterization of the mechanical connection.

the desired properties. Note also that $\alpha_\mu(q)$ lies in the orthogonal space to $T_q^*Q \cap \mathbf{J}^{-1}(\mu)$ in the fiber T_q^*Q relative to the bundle metric on T^*Q defined by the Riemannian metric on Q. It also follows that $\alpha_\mu(q)$ is the unique critical point of the kinetic energy of the bundle metric on T^*Q restricted to the fiber $T_q^*Q \cap \mathbf{J}^{-1}(\mu)$.

3.13 Coupled Planar Rigid Bodies

In §3.13 of the text we studied some of the basic properties of the system of coupled rigid bodies. Here we give some additional information on this system and its extension to multibody problems.[2]

Equilibria and Stability by the Energy-Casimir Method. We now use Arnold's energy-Casimir method, to determine the equilibrium points and their stability for the system of two coupled rigid bodies. An equivalent alternative to this method is to look for critical points of H given by (3.13.42) in the text, which is an expression in (θ, ν)-space and then to test $\delta^2 H$ for definiteness at these equilibria.

To search for equilibria, we can look directly at Hamilton's equations on P. The conditions $\dot{\mu}_1 = \dot{\mu}_2 = 0$ give

$$\frac{\partial H}{\partial \theta} = 0; \tag{3.13.45}$$

that is,

$$-\frac{1}{2}(\mu_1, \mu_2)\mathbf{J}^{-1}\frac{\partial \mathbf{J}}{\partial \theta}\mathbf{J}^{-1}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = 0. \tag{3.13.46}$$

Clearly,

$$\frac{d\mathbf{J}}{d\theta} = \begin{pmatrix} 0 & \varepsilon\lambda' \\ \varepsilon\lambda' & 0 \end{pmatrix}, \tag{3.13.47}$$

---

[2]This section is based on Sreenath, Oh, Krishnaprasad, and Marsden [1988] and on Oh, Sreenath, Krishnaprasad, and Marsden [1989].

so we get

$$-\frac{1}{2}(\omega_1, \omega_2)\begin{pmatrix} 0 & \varepsilon\lambda' \\ \varepsilon\lambda' & 0 \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = 0; \tag{3.13.48}$$

that is,

$$-\varepsilon\lambda'\omega_1\omega_2 = 0. \tag{3.13.49}$$

The equilibrium condition $\dot{\theta} = 0$ becomes $\tilde{I}_1\mu_1 - \varepsilon\lambda\mu_2 = \tilde{I}_2\mu_2 - \varepsilon\lambda\mu_1$ or, equivalently, $\omega_1 = \omega_2$. Thus, the equilibria are given by

(i) $\omega_1 = \omega_2 = 0$, or

(ii) $\omega_1 = \omega_2 \neq 0, \lambda' = 0$.

For simplicity, choose the reference configuration so that $\mathbf{d}_{12}$ and $\mathbf{d}_{21}$ are parallel. Then

$$\lambda'(\theta) = \mathbf{d}_{12} \cdot \mathbf{d}_{21} \sin\theta$$

so the equilibria in case (ii) occur when

(ii′) either (a) $\mathbf{d}_{12} = 0$ or $\mathbf{d}_{21} = 0$, or (b) $\theta = 0$ or π. The case in which $\theta = \pi$ corresponds to the case of folded bodies, while $\theta = 0$ corresponds to extended (stretched out) bodies.

The first step in the energy-Casimir method is to realize the equilibria as critical points of $H + C$. We calculate that

$$\begin{aligned} \frac{\partial H}{\partial \theta} &= \varepsilon\lambda'\omega_1\omega_2 \\ \frac{\partial H}{\partial \mu_1} = \omega_1; &\quad \frac{\partial H}{\partial \mu_2} - \omega_2, \end{aligned} \tag{3.13.50}$$

where

$$\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \mathbf{J}^{-1}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} \tilde{I}_2\mu_1 - \varepsilon\lambda\mu_2 \\ \tilde{I}_1\mu_2 - \varepsilon\lambda\mu_1 \end{pmatrix}.$$

The first variation is

$$\mathbf{d}(H + C) = \frac{\partial H}{\partial \theta}d\theta + \left(\frac{\partial H}{\partial \mu_1} + \Phi'\right)d\mu_1 + \left(\frac{\partial H}{\partial \mu_2} + \Phi'\right)d\mu_2, \tag{3.13.51}$$

from which we see that critical points of $H + C$ correspond to equilibria provided

$$\Phi'(\mu_e) = -\left(\frac{\partial H}{\partial \mu_1}\right)_e = -\left(\frac{\partial H}{\partial \mu_1}\right)_e, \tag{3.13.52}$$

where the subscript e means evaluation at the equilibrium. Here, $\Phi''(\mu_e)$ is arbitrary.

The matrix of the second variation of $H + C$ at equilibrium is

$$\delta^2(H+C) = \begin{pmatrix} \dfrac{\partial^2 H}{\partial \theta^2} & \dfrac{\partial^2 H}{\partial \theta \partial \mu_1} & \dfrac{\partial^2 H}{\partial \theta \partial \mu_2} \\ \dfrac{\partial^2 H}{\partial \theta \partial \mu_1} & \dfrac{\partial^2 H}{\partial \mu_1^2} + \Phi'' & \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} + \Phi'' \\ \dfrac{\partial^2 H}{\partial \theta \partial \mu_2} & \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} + \Phi'' & \dfrac{\partial^2 H}{\partial \mu_2^2} + \Phi'' \end{pmatrix}, \qquad (3.13.53)$$

where

$$\begin{pmatrix} \dfrac{\partial^2 H}{\partial \mu_1^2} & \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} \\ \dfrac{\partial^2 H}{\partial \mu_1 \partial \mu_2} & \dfrac{\partial^2 H}{\partial \mu_2^2} \end{pmatrix} = \mathbf{J}^{-1} = \frac{1}{\Delta} \begin{pmatrix} \tilde{I}_2 & -\varepsilon\lambda \\ -\varepsilon\lambda & \tilde{I}_1 \end{pmatrix},$$

$$\frac{\partial^2 H}{\partial \theta \partial \mu_1} = -\frac{\varepsilon\lambda'}{\Delta^2}(\tilde{I}_2\omega_2 - \varepsilon\lambda\omega_1), \qquad \frac{\partial^2 H}{\partial \theta \partial \mu_2} = -\frac{\varepsilon\lambda'}{\Delta^2}(-\varepsilon\lambda\omega_2 + \tilde{I}_1\omega_1),$$

and

$$\frac{\partial^2 H}{\partial \theta^2} = \frac{\partial}{\partial \theta}\left[-\varepsilon\lambda' \frac{\partial H}{\partial \mu_1}\frac{\partial H}{\partial \mu_2}\right] = -\varepsilon\lambda''\omega_1\omega_2 - \varepsilon\lambda' \frac{\partial^2 H}{\partial \theta \partial \mu_1}\omega_2 - \varepsilon\lambda'\omega_1 \frac{\partial^2 H}{\partial \theta \partial \mu_2}.$$

At equilibrium, $\lambda = \pm d_1 d_2$ ($+$ if $\theta = 0, -$ if $\theta = \pi$) and so

$$\mathbf{J}^{-1} \frac{1}{(\tilde{I}_1\tilde{I}_2 - \varepsilon^2 d_1^2 d_2^2)} \begin{bmatrix} \tilde{I}_2 & \mp\varepsilon d_1 d_2 \\ \mp\varepsilon d_1 d_2 & \tilde{I}_1 \end{bmatrix},$$

$$\frac{\partial^2 H}{\partial \theta \partial \mu_1} = 0 = \frac{\partial^2}{\partial \theta \partial \mu_2}, \quad \text{and} \quad \frac{\partial^2 H}{\partial \theta^2} = -\varepsilon\lambda''\omega_e^2 = \pm\varepsilon d_1 d_2 \omega_e^2,$$

where $\omega_e = \omega_1 = \omega_2 \neq 0$ at equilibrium. Thus we get

$$\delta^2(H+C) = \begin{pmatrix} \pm\varepsilon\lambda d_1 d_2 \omega_e^2 & 0 \\ 0 & \mathbf{J}^{-1} + \Phi'' \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \end{pmatrix}. \qquad (3.13.54)$$

This matrix is clearly positive definite if $d_1 \neq 0, d_2 \neq 0$ if $\theta = 0$ ($+$ sign) and $\Phi''(\mu_e) \geq 0$ and is indefinite for any choice of $\Phi''(\mu_e)$ if $\theta = \pi$.

Another way to do the stability analysis is to use the reduced Hamiltonian on T^*S^1. After completing squares, H will have the form of kinetic plus potential energy with effective potential given by

$$V(\theta) = \frac{1}{2\Delta}\left[\frac{1}{4}\mu^2(\tilde{I}_1 + \tilde{I}_2 - 2\varepsilon\lambda) + \frac{(\tilde{I}_1 + \tilde{I}_2)^2\mu^2}{4(\tilde{I}_1 + \tilde{I}_2 + 2\varepsilon\lambda)}\right]. \qquad (3.13.55)$$

Minima of V are then the stable equilibria while maxima are unstable. The following theorem summarizes the situation.

3.13.2 Theorem. *The dynamics of the* 2*-body problem is completely integrable and contains one stable relative equilibrium solution* ($\theta = 0$*—the stretched-out case) and one unstable relative equilibrium solution* ($\theta = \pi$*—the folded-over case). The reduced dynamics contains a homoclinic orbit joining the unstable equilibrium to itself.*

For three or more bodies, this method of looking for minima of the potential will not work in a naive way because the symplectic structures on the symplectic leaves will have magnetic terms. The general theory for dealing with this situation is given in Simo, Lewis, and Marsden [1991]. See the internet supplement for further details on the case of multibody problems.

Multibody Problems. The Hamiltonian formulation of that section extends to systems of N planar rigid bodies connected to form a ***tree structure***. Since the general statement of this result requires significant additional notation we limit ourselves to the special case of a ***chain*** of N bodies.

3.13.3 Theorem. *The total kinetic energy (Hamiltonian) for an open chain of* N *planar rigid bodies connected together by hinge joints has the form*

$$H = \mu^T \cdot \mathbf{J}^{-1} \cdot \mu \tag{3.13.56}$$

where $\mu = (\mu_1, \mu_2, \ldots, \mu_N)^T$ *is the momentum vector and* $\mathbf{J}$ *is the corresponding* $N \times N$ *inertia matrix which is a function of the set of relative (or joint) angles between adjacent bodies. The reduced dynamics takes the form*

$$\begin{aligned}
\dot{\mu}_1 &= \frac{\partial H}{\partial \theta_{2,1}} \\
\dot{\mu}_2 &= \frac{\partial H}{\partial \theta_{3,2}} - \frac{\partial H}{\partial \theta_{2,1}} \\
\dot{\mu}_i &= \frac{\partial H}{\partial \theta_{i+1,i}} - \frac{\partial H}{\partial \theta_{i,i-1}} \\
\dot{\mu}_N &= -\frac{\partial H}{\partial \theta_{N,N-1}} \\
\dot{\theta}_{i+1,i} &= \frac{\partial H}{\partial \theta_{i+1}} - \frac{\partial H}{\partial \mu_i} \quad \textit{for} \quad i = 1, \ldots, N-1)
\end{aligned} \tag{3.13.57}$$

where $\theta_{i+1,i}$ *is the joint angle between body* $i+1$ *and body* i*. The associated Poisson structure is given by*

$$\{f, g\} = \sum_{i=0}^{N-1} \left(\frac{\partial f}{\partial \mu_i} - \frac{\partial f}{\partial \mu_{i+1}} \right) \frac{\partial g}{\partial \theta_{i+1,i}} - \frac{\partial f}{\partial \theta_{i+1,i}} \left(\frac{\partial g}{\partial \mu_i} - \frac{\partial g}{\partial \mu_{i+1}} \right). \tag{3.13.58}$$

This is proven in a way similar to the two-body case. The structure of equilibria and the associated stability analysis become quite complex and interesting as the number of interconnected bodies increases. A mixture of topological and geometric methods may be necessary to extract useful information on the phase portraits.

In the remainder of this section, we illustrate some of the complexities of multibody problems by discussing of the equilibria and stability for a system of three planar rigid bodies connected by hinge joints (see figure 3.13.1).

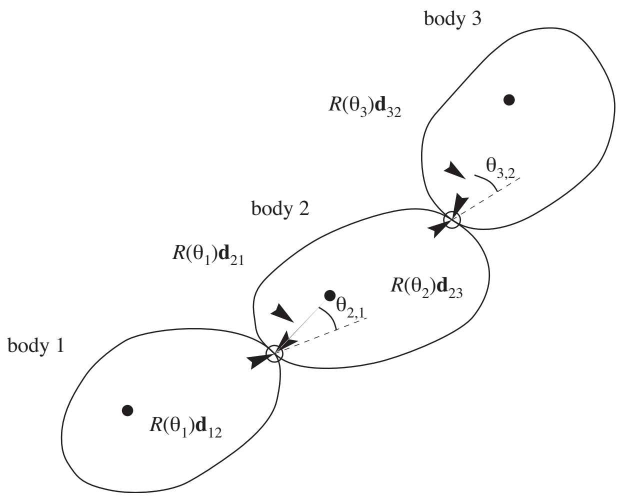

FIGURE 3.13.1. Planar three-body system.

The Hamiltonian of the planar three-body problem is given by equation (3.13.56) with the momentum vector μ and the coefficient of inertia matrix $\mathbf{J}$ being defined as follows:

$$\mu = (\mu_1, \mu_2, \mu_3)^T,$$

$$\mathbf{J} = \begin{pmatrix} \tilde{I}_1 & \tilde{\lambda}_{12}(\theta_{2,1}) & \tilde{\lambda}_{31}(\theta_{2,1} + \theta_{3,2}) \\ \tilde{\lambda}_{12}(\theta_{2,1}) & \tilde{I}_1 & \tilde{\lambda}_{23}(\theta_{3,2}) \\ \tilde{\lambda}_{31}(\theta_{2,1} + \theta_{3,2}) & \tilde{\lambda}_{23}(\theta_{3,2}) & \tilde{I}_3 \end{pmatrix}. \tag{3.13.59}$$

Here $\theta_{2,1}$ and $\theta_{3,2}$ are the relative angles between bodies 2 and 1, and bodies 3 and 2, respectively. The coefficients of inertia $\tilde{I}_i$ and $\tilde{\lambda}_{ij}$ are given by

$$\begin{aligned}
\tilde{I}_1 &= [I_1 + (\varepsilon_{12}\varepsilon_{31}) \langle \mathbf{d}_{12}, \mathbf{d}_{12} \rangle], \\
\tilde{I}_2 &= [I_2 + \varepsilon_{12} \langle \mathbf{d}_{21}, \mathbf{d}_{21} \rangle + \varepsilon_{23} \langle \mathbf{d}_{23} - \mathbf{d}_{23} \rangle \\
&\quad + \varepsilon_{31} \langle (\mathbf{d}_{23} - \mathbf{d}_{21}), (\mathbf{d}_{23} - \mathbf{d}_{21}) \rangle] \\
\tilde{I}_3 &= [I_3 + (\varepsilon_{23} + \varepsilon_{31}) \langle \mathbf{d}_{32}, \mathbf{d}_{32} \rangle] \\
\tilde{\lambda}_{12}(\theta_{2,1}) &= [\varepsilon_{12} \lambda_{(-\mathbf{d}_{21}, \mathbf{d}_{12})}(\theta_{2,1}) + \varepsilon_{31} \lambda_{(\mathbf{d}_{23}, -\mathbf{d}_{21}, \mathbf{d}_{12}))}(\theta_{2,1})] \\
\tilde{\lambda}_{23}(\theta_{3,2}) &= [\varepsilon_{23} \lambda_{(-\mathbf{d}_{32}, \mathbf{d}_{23})}(\theta_{3,2}) + \varepsilon_{31} \lambda_{(-\mathbf{d}_{32}, \mathbf{d}_{23}, \mathbf{d}_{21}))}(\theta_{3,2})] \\
\tilde{\lambda}_{31}(\theta_{2,1}) + \theta_{3,2}) &= \varepsilon_{31} \lambda_{(\mathbf{d}_{32}, \mathbf{d}_{12})}(\theta_{2,1} + \theta_{3,2}) \\
\varepsilon_{ij} &= \frac{m_i m_j}{m_1 + m_2 + m_3}, \quad i \neq j \quad \text{and} \quad i, j = 1, 2, 3 \\
\lambda_{(x,y)}(\alpha) &= \mathbf{x} \cdot \mathbf{y} \cos \alpha + [\mathbf{x} \times \mathbf{y}] \sin \alpha,
\end{aligned}$$

where the m_i and I_i are the mass and inertia respectively of the body i, and the $\mathbf{d}_{ij}$ are defined as in the figure.

The dynamics of a three-body system of planar, rigid bodies in the Hamiltonian setting is given by:

$$\begin{aligned}
\dot{\mu}_1 &= \frac{\partial H}{\partial \theta_{2,1}} \\
\dot{\mu}_2 &= -\frac{\partial H}{\partial \theta_{2,1}} + \frac{\partial H}{\partial \theta_{3,2}} \\
\dot{\mu}_3 &= -\frac{\partial H}{\partial \theta_{3,2}} \\
\dot{\theta}_{2,1} &= \frac{\partial H}{\partial \mu_2} - \frac{\partial H}{\partial \mu_1} \\
\dot{\theta}_{3,2} &= \frac{\partial H}{\partial \mu_3} - \frac{\partial H}{\partial \mu_2}
\end{aligned} \tag{3.13.60}$$

Using Equation 3.13.60 note that the sum $\mu_1 + \mu_2 + \mu_3$ of momentum variables is constant in time.

In Sreenath, Oh, Krishnaprasad, and Marsden [1988] and Oh [1987], it is shown that for three coupled rigid bodies there are either 4 or 6 relative equilibria and the bifurcations between these are determined as a function of the system parameters. It is also shown that near the stable stretched out relative equilibrium, there are relative periodic orbits distinguished by symmetry type. This is done using the Montaldi, Roberts, and Stewart [1988] symmetric version of the Moser-Weinstein theorem (Weinstein [1973, 1978a] and Moser [1976]). Also, it is shown that the dynamics is, in general, not integrable by using the Poincaré–Melnikov method.

4

Introduction to Aspects of Geometric Control Theory

4.7 An Alternative View of Hamiltonian and Lagrangian Control Systems

This supplement presents an alternative point of view from that discussed in Section 4.6 of the text for defining Hamiltonian (or Lagrangian) systems with external forces, such as controls. More precisely, the section discusses briefly how one might characterize Hamiltonian control systems among a suitable class of general nonlinear control systems with inputs and outputs. This is an approach which is developed in detail in the book of Crouch and van der Schaft [1987]. This generalizes the work of Brockett and Rahimi [1972] which characterizes linear Hamiltonian systems among the set of linear input-output maps. The approach here is variational.

One generalizes the definition of a Hamiltonian vector field as a parameterization of a Lagrangian submanifold, as discussed in section 3.3. If in system

$$\begin{aligned} \dot{x} &= f(x,u), \\ y &= h(x,u), \end{aligned} \tag{4.7.1}$$

$N = P = \mathbb{R}^k$, set

$$\omega^e = \sum_{i=1}^{k} du_i \wedge dy_i$$

as a symplectic form on $\mathbb{R}^{2k}$, parameterized by coordinates $(u_1, \ldots, u_k, y_1, \ldots, y_k)$. Assume that M is symplectic with symplectic form ω, then $TM \oplus$

$\mathbb{R}^{2k}$ is a symplectic manifold with symplectic form $\dot{\omega} + \omega^e$. (The notation $\dot{\omega}$ was defined in §3.7).

If $H(x, u)$ is a function on $M \oplus \mathbb{R}^k \subset TM \oplus \mathbb{R}^{2k}$, we may consider the following system generated by H

$$\begin{aligned} \dot{x} &= X_{H(x,u)}(x) \\ y &= \frac{\partial H}{\partial u}(x, u) \end{aligned} \tag{4.7.2}$$

where X_H is the Hamiltonian vector field on M (with symplectic form ω). System (4.7.2) defines a Lagrangian submanifold N of $TM \oplus \mathbb{R}^{2k}$ with symplectic form $\dot{\omega} + \omega^e$, $N = \{(x, u), (X_H, (\partial H/\partial u))\}$. (Lagrangian submanifolds were defined in Section 3.3). It follows that one may define a Hamiltonian control system (4.6.4), with $N = P = \mathbb{R}^k$, as globally parameterized Lagrangian submanifolds of $TM \oplus \mathbb{R}^{2k}$ with symplectic form $\dot{\omega} + \omega^e$. See van der Schaft [1982, 1983] for a detailed discussion.

If H is affine in the coordinates u_i, the vector field X_H is also affine in u_i, and so one can directly compare X_H with X_L (and $L_{F*}X_L = X_H$) in (3.6.0), where $M = TQ$ and $F(t)$ has the form

$$F(t) = \sum_{i=1}^{k} \eta_i u_i(t), \qquad \eta_i \in \Gamma(T^*Q). \tag{4.7.3}$$

Unfortunately, the restrictions imposed by the structure (4.7.2) do not make for direct comparisons, except in some special cases. However, the model (4.7.2) has some added attractions when studying the input-output properties of systems (4.7.1).

In the work of Brockett and Rahimi [1972] , it was shown that one could characterize initialized, linear, time invariant, Hamiltonian systems (4.7.2) with input-output maps

$$y(t) = \int_0^t W(t - \sigma) u(\sigma) d\sigma \tag{4.7.4}$$

by the additional symmetry property of the impulse response

$$W(t) = -W(-t)^T. \tag{4.7.5}$$

This result was generalized to nonlinear systems (4.7.2) by Crouch and van der Schaft [1987] , using a variational property. Specifically, if (u, y) is the input response of a system (4.7.1), with $N = V = \mathbb{R}^k$, $P = V^* = \mathbb{R}^k$, and defined on the infinite interval $(-\infty, \infty)$, then we define an admissible variation of (u, y) to be a mapping

$$(t, \epsilon) \mapsto (u(t, \epsilon), y(t, \epsilon)), (t, \epsilon) \in (-\infty, \infty) \times (-\delta, \delta)$$

satisfying

(i) $(u(t, 0), y(t, 0)) = (u(t), y(t))$

(ii) $(u(t, \epsilon), y(t, \epsilon))$ is an input-output response of a given system (4.7.1) for each $\epsilon \in (-\delta, \delta)$

(iii) $(\delta u(t), \delta y(t)) = (\partial/\partial\epsilon)|_{\epsilon=0}(u(t, \epsilon), y(t, \epsilon))$ has compact support along (u, y).

We define a manifold $\mathcal{M}_V$ of maps consisting of all input-output behaviors (u, y) of systems (1), corresponding to the same input-output space $N = V = \mathbb{R}^k$, $P = V^* = \mathbb{R}^{k^*}$. The set of all admissible variations of $(u, y) \in \mathcal{M}_V$ defines a set of variational fields $(\delta u, \delta y)$ along (u, y), which may be interpreted as the tangent space to $\mathcal{M}_V$ at (u, y). There is a skew form on this tangent space defined by

$$\omega_{(u,y)}((\delta_1 u, \delta_1 y), (\delta_2 u, \delta_2 y)) = \int_{-\infty}^{\infty} (\delta_2 y(\delta_1 u) - \delta_1 y(\delta_2 u))dt. \qquad (4.7.6)$$

The notion of isotropic, co-isotropic and Lagrangian submanifolds can be defined as it is for finite dimensional manifolds. It is shown in Crouch and van der Schaft [1987] that the set of input-output behaviors for Hamiltonian systems (4.7.2) form a Lagrangian submanifold of the manifold $\mathcal{M}_V$.

It is interesting to note that this variational characterization of Hamiltonian systems is related to the classical inverse problem in mechanics, (see, for example Anderson and Duchamp [1980], Santilli [1978], and references therein.) In its most basic version, this problem considers an implicit system of second order equations

$$R_i(q, \dot{q}, \ddot{q}) = 0, \quad 1 \le i \le n, \quad q \in \mathbb{R}^n,$$

where $[\partial^2 R_k/\partial\ddot{q}_i\partial\ddot{q}_j]_{1\le i,j\le n}$ is an invertible matrix for all values of $(q, \dot{q}, \ddot{q})$, and asks: "when this system is equivalent to a system of Lagrangian equations for some choice of Lagrangian function L?" This problem is in fact a special case of the problem of determining when there exists a function $H(x, u)$ such that system (4.7.1) coincides with the Hamiltonian system (4.7.2). This problem is answered in detail in Crouch and van der Schaft [1987]. Clearly the conditions that are developed in this work ensure that the symplectic form (4.7.6) vanishes and, for linear systems, they are equivalent to the condition (4.7.5) of Brockett and Rahimi [1972]

We shall return to systems of this type. Good discussions are given in Krishnaprasad [1985] and Sanchez [1989].

5

Nonholonomic Mechanics

5.2 The Lagrange–d'Alembert Principle

The Equations of Motion in terms of the Constrained Lagrangian. Here we give some of the calculations necessary to derive the equations of motion for a nonholonomic system in terms of the constrained Lagrangian that were stated in §5.2 of the text, namely as

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot{r}^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A_\alpha^a \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot{s}^b} B_{\alpha\beta}^b \dot{r}^\beta, \tag{5.2.17}$$

where

$$B_{\alpha\beta}^b = \left(\frac{\partial A_\alpha^b}{\partial r^\beta} - \frac{\partial A_\beta^b}{\partial r^\alpha} + A_\alpha^a \frac{\partial A_\beta^b}{\partial s^a} - A_\beta^a \frac{\partial A_\alpha^b}{\partial s^a} \right). \tag{5.2.18}$$

The derivation of these equations of motion proceeds as follows: using the relationships

$$\frac{\partial L_c}{\partial \dot{r}^\alpha} = \frac{\partial L}{\partial \dot{r}^\alpha} - A_\alpha^b \frac{\partial L}{\partial \dot{s}^b}$$
$$\frac{\partial L_c}{\partial r^\alpha} = \frac{\partial L}{\partial r^\alpha} - \frac{\partial L}{\partial \dot{s}^b} \left(\frac{\partial A_\beta^b}{\partial r^\alpha} \dot{r}^\beta \right)$$
$$\frac{\partial L_c}{\partial s^a} = \frac{\partial L}{\partial s^a} - \frac{\partial L}{\partial \dot{s}^b} \left(\frac{\partial A_\beta^b}{\partial s^a} \dot{r}^\beta \right)$$

and substituting L_c into the nonholonomic Lagrange equations yields

$$\begin{aligned}
&\frac{d}{dt}\frac{\partial L_c}{\partial \dot r^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} \\
&\qquad = \left(\frac{d}{dt}\frac{\partial L}{\partial \dot r^\alpha} - \frac{\partial L}{\partial r^\alpha}\right) - A^a_\alpha\left(\frac{d}{dt}\frac{\partial L}{\partial \dot s^a} - \frac{\partial L}{\partial s^a}\right) \\
&\qquad\quad - \frac{\partial L}{\partial \dot s^b}\frac{d}{dt}A^b_\alpha + \frac{\partial L}{\partial \dot s^b}\frac{\partial A^b_\beta}{\partial r^\alpha}\dot r^\beta - A^a_\alpha \frac{\partial L}{\partial \dot s^b}\frac{\partial A^b_\beta}{\partial s^a}\dot r^\beta \\
&\qquad = \left(\frac{d}{dt}\frac{\partial L}{\partial \dot r^\alpha} - \frac{\partial L}{\partial r^\alpha}\right) - A^a_\alpha\left(\frac{d}{dt}\frac{\partial L}{\partial \dot s^a} - \frac{\partial L}{\partial s^a}\right) \\
&\qquad\quad + \frac{\partial L}{\partial \dot s^b}\left(\frac{\partial A^b_\beta}{\partial r^\alpha} - \frac{\partial A^b_\alpha}{\partial r^\beta} + A^a_\beta \frac{\partial A^b_\alpha}{\partial s^a} - A^a_\alpha \frac{\partial A^b_\beta}{\partial s^a}\right)\dot r^\beta.
\end{aligned}$$

Hence the equations of motion can be written as claimed.

Note that L_c is a degenerate Lagrangian in the sense that it does not depend on $\dot s$. Also note that thinking of s as a cyclic variable does not lead to conservation laws in the usual way because of the constraints.

To see how the right hand side of the constrained Lagrange d'Alembert equation is related to the curvature of the Ehresmann connection of $A = \omega^a(\partial/\partial s^a)$, let $d\omega^b$ be the exterior derivative of ω^b:

$$\begin{aligned}
d\omega^b &= d(ds^b + A^b_\alpha dr^\alpha) \\
&= \frac{\partial A^b_\alpha}{\partial r^\beta} dr^\beta \wedge dr^\alpha - \frac{\partial A^b_\alpha}{\partial s^a} A^a_\beta dr^\beta \wedge dr^\alpha. \qquad (5.2.19)
\end{aligned}$$

Contracting $d\omega^b$ with $\dot q$ yields

$$\begin{aligned}
d\omega^b(\dot q, \cdot) &= \frac{\partial A^b_\alpha}{\partial r^\beta}\dot r^\beta dr^\alpha - \frac{\partial A^b_\alpha}{\partial s^a} A^a_\beta \dot r^\beta dr^\alpha - \frac{\partial A^b_\alpha}{\partial r^\beta}\dot r^\alpha dr^\beta + \frac{\partial A^b_\alpha}{\partial s^a} A^a_\beta \dot r^\alpha dr^\beta \\
&= \left(\frac{\partial A^b_\alpha}{\partial r^\beta} + \frac{\partial A^b_\beta}{\partial s^a}A^a_\alpha - \frac{\partial A^b_\beta}{\partial r^\alpha} - \frac{\partial A^b_\alpha}{\partial s^a}A^a_\beta\right)\dot r^\beta dr^\alpha \\
&= B^b_{\alpha\beta}\dot r^\alpha dr^\beta. \qquad (5.2.20)
\end{aligned}$$

Combining all of these calculations, we can write the equations of motion for the constrained system as

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot r^\alpha} - \frac{\partial L_c}{\partial r^\alpha} + A^a_\alpha \frac{\partial L_c}{\partial s^a} = -\frac{\partial L}{\partial \dot s^a}\mathbf{d}\omega^a\left(\dot q, \frac{\partial}{\partial r^\alpha}\right). \qquad (5.2.21)$$

The left-hand side of (5.2.21) may be checked to be the variational derivative of the constrained Lagrangian. The right-hand side consists of the forces that maintain the constraints. In the special case that the constraints are holonomic, $\mathbf{d}\omega^a = 0$ since $\mathbf{d}\omega^a$ represents the curvature and the curvature measures the lack of integrability of the constraints; when they are integrable, we have, by definition, the holonomic case. In this case, equation (5.2.21) reduces to the usual form of Lagrange's equations. This verifies that for holonomic systems it is appropriate to "plug in the constraints" before applying Lagrange's equations.

5.3 An Invariant Approach to Nonholonomic Mechanics

In this section we consider an invariant approach to Lagrangian mechanics in general and nonholonomic mechanics in particular. We follow here the approach of Vershik [1984] (see also Wang and Krishnaprasad [1992]). A general invariant approach to Lagrangian mechanics is also discussed for example in Marsden and Ratiu [1999].

Lagrangian Mechanics. We consider firstly a general invariant formulation of Lagrangian mechanics without constraints.

Let TQ be the tangent bundle of Q, an n-dimensional manifold, with its canonical projection and let it be locally co-ordinitized by $(q, \dot{q}) \sim (q, v)$. Here q, v represent n-vectors. We consider here notions of verticality and horizontality with respect to the trivial connection on TQ, i.e., combinations of vectors $\frac{\partial}{\partial q^i}$ are horizontal and combinations of vectors $\frac{\partial}{\partial v^i}$ are vertical. (It is important to bear this in mind when discussing constraints below, so as not to confuse this discussion with the connection given by the constraints discussed elsewhere.)

Since Lagrangian equations are second order we also need the second tangent bundle $T(TQ)$ with projection $T_{\tau_Q} : T(TQ) \to TQ$ locally coordinitized by $(q, v, v, \dot{v})$. (A useful refereerence in this regard is Marsden, Patrick, and Shkoller [1998].) In fact we want vector fields of the type

$$\ddot{q} = f(q, \dot{q}) \tag{5.3.1}$$

or

$$\begin{aligned} \dot{q} &= v \\ \dot{v} &= f(q, v). \end{aligned} \tag{5.3.2}$$

Written as a vector field a second order equation takes the form

$$X_{q,v} = X_s = v \, \frac{\partial}{\partial q} + f \, \frac{\partial}{\partial v} \,, \tag{5.3.3}$$

(with the obvious summation notation).

Invariantly we may write this as $T_{\tau_Q} X_s = v$—its projection onto the tangent space to Q is just v. We call such a vector field X_s a "special" vector field.

Now one can formulate Lagrangian mechanics in an invariant fashion by defining fields on TQ as follows: Let T_qQ and $T_{q,v}TQ$ define the tangent space to Q at q and TQ at (q,v) respectively. We set up a canonical map from T_qQ to $T_{q,v}TQ$ which takes the fiber over q (locally vectors of the form $a\frac{\partial}{\partial q}$) into the fiber over $T_{q,v}TQ$ (vectors of the form $a\frac{\partial}{\partial q} + b\frac{\partial}{\partial v}$). Locally this map is defined by

$$\gamma_{q,v}\left(a\frac{\partial}{\partial q}\right) \;=\; a\frac{\partial}{\partial v}. \tag{5.3.4}$$

Define a tensor field on TQ, called the ***principal tensor field***, by

$$P_{q,v} = \gamma_{q,v}\; T_{\tau_{q,v}}. \tag{5.3.5}$$

So locally we have

$$\begin{aligned} P_{q,v}\left(a\frac{\partial}{\partial q} + b\frac{\partial}{\partial v}\right) &= \gamma_{q,v}\; T_{\tau_{q,v}}\left(a\frac{\partial}{\partial q} + b\frac{\partial}{\partial v}\right) \\ &= \gamma_{q,v}\left(a\frac{\partial}{\partial q}\right) = a\frac{\partial}{\partial v}. \end{aligned} \tag{5.3.6}$$

The dual tensor field P^* acts locally on forms as follows

$$P^*(adq + bdv) = bdq. \tag{5.3.7}$$

This can be easily checked: we have

$$\begin{aligned} &\left\langle P^*(adq + bdv), c\frac{\partial}{\partial q} + d\frac{\partial}{\partial v}\right\rangle \\ &= \left\langle adq + bdv, P\left(c\frac{\partial}{\partial q} + d\frac{\partial}{\partial v}\right)\right\rangle \\ &= \left\langle adq + bdv, c\frac{\partial}{\partial v}\right\rangle = bc. \end{aligned} \tag{5.3.8}$$

Hence

$$P^*(adq + bdv) = bdq. \tag{5.3.9}$$

Now we define another key concept—the *fundamental* vector field on TQ: this is a field with coordinates

$$\Phi_{q,v} = \gamma_{q,v} v\frac{\partial}{\partial q} = v^i\ \frac{\partial}{\partial v^i}.$$

Clearly a vector field X is special if and only if $PX = \Phi$, since in local coordinates a special vector field is of the form

$$X_{q,v} \;= v^i\frac{\partial}{\partial q^i} + \cdots\ . \tag{5.3.10}$$

Now let L be the Lagrangian—as usual a smooth function on TQ. We note that locally

$$P^*(dL) = P^*\left(\frac{\partial L}{\partial q^i}dq^i + \frac{\partial L}{\partial \dot{q}^i}\,d\dot{q}^i\right) = \frac{\partial L}{\partial \dot{q}^i}\,dq^i \tag{5.3.11}$$

—clearly a horizontal 1-form since it annihilates vertical vector fields $a\frac{\partial}{\partial v}$ (Vershik calls this form the impulse field of the Lagrangian.) One can identify horizontal 1-forms in Lagrangian mechanics as forces—see later).

We define the Lagrangian 2-form to be

$$\begin{aligned}\Omega_L &= -d(P^*dL)\\ &= -d\left(\frac{\partial L}{\partial \dot{q}^i}\,dq^i\right)\\ &= -\frac{\partial^2 L}{\partial q^j \partial v^i}\,dq^j \wedge dq^i + \frac{\partial^2 L}{\partial v^j \partial v^i}\,dq^j \wedge dv^i.\end{aligned} \tag{5.3.12}$$

Recalling the Legendre transformation $p_i = \frac{\partial L}{\partial \dot{q}^i}$ we see that the image of Ω_L under the Legendre transformation is the canonical symplectic form on the cotangent bundle. Similarly the Hamiltonian (energy) is

$$H_L = dL(\Phi) - L. \tag{5.3.13}$$

Note that in local coordinates

$$dL(\Phi) = dL\left(v^i\frac{\partial}{\partial v^i}\right) = \frac{\partial L}{\partial v^i}v^i = p_i v^i\,, \tag{5.3.14}$$

(with the summation convention).

Now we can formulate the Lagrange d'Alembert principle as follows:

5.3.1 Definition (The Lagrange D'Alembert Principle)**.** *The vector field Y describing the mechanical trajectories of motion is given by*

$$\Omega_L(X, Y) = dH_L(Y) + \omega(Y) \tag{5.3.15}$$

where ω is the 1-form describing the exterior forces and X is a special vector field.

In the absence of exterior forces we recover the Lagrange equations as follows. Let

$$X = v\frac{\partial}{\partial q} + \dot{v}\frac{\partial}{\partial v}\,, \qquad Y = a\frac{\partial}{\partial q} + b\frac{\partial}{\partial v}. \tag{5.3.16}$$

Then

$$\Omega_L(X,Y) = -\frac{\partial^2 L}{\partial q^j \partial v^i}(v^j a^i - a^j v^i) + \frac{\partial^2 L}{\partial v^j \partial v^i}(v^j b^i - a^j \dot{v}^i)\,, \qquad (5.3.17)$$

$$\begin{aligned} dH_L(Y) &= d\ \left(\frac{\partial L}{\partial v^i}v^i - L\right)(Y) \qquad (5.3.18)\\ &= \Big(\frac{\partial L}{\partial v^i}dv^i + \frac{\partial^2 L}{\partial q^j \partial v^i}v^i dq^j + \frac{\partial^2 L}{\partial v^i \partial v^j}v^i dv^j \\ &\quad - \frac{\partial L}{\partial q^i}dq^i - \frac{\partial L}{\partial v^i}dv^i\Big)(Y)\\ &= \frac{\partial L}{\partial v^i}b^i + \frac{\partial^2 L}{\partial q^j \partial v^i}v^i a^j + \frac{\partial^2 L}{\partial v^j \partial v^i}v^i b^j - \frac{\partial L}{\partial q^i}a^i - \frac{\partial L}{\partial v^i}\ b^i\,. \end{aligned}$$

Equating coefficients of a^i and b^i which are arbitrary we get:

$$\begin{aligned} -\frac{\partial^2 L}{\partial q^j \partial v^i}\ v^j + \frac{\partial^2 L}{\partial q^i \partial v^j}\ v^j - \frac{\partial^2 L}{\partial v^j \partial v^i}\ \dot{v}^j &= -\frac{\partial L}{\partial q^i} + \frac{\partial^2 L}{\partial q^i \partial v^j}\ v^j \\ \frac{\partial^2 L}{\partial v^i \partial v^j}\ v^j &= \frac{\partial L}{\partial v^i} + \frac{\partial^2 L}{\partial v^i \partial v^j}\ v^j - \frac{\partial L}{\partial v^i} \end{aligned}$$

The second equation is an identity while the first equation is just

$$\frac{d}{dt}\ \left(\frac{\partial L}{\partial v^i}\right) = \frac{\partial L}{\partial q^i}. \qquad (5.3.19)$$

as required.

Constrained Dynamics. We now consider the invariant formulation of dynamics with constraints.

As in Vershik [1984], we assume a slightly more general form of the constraints—we assume that they define a distribution on TQ, i.e., they define at each point a subspace of $TTQ(q, \dot{q})$. This fits naturally with the definition of 2nd order systems, although it is more general than needed. Hence we can define the constraints to be 1-forms on TQ (a codistribution of TQ) and they thus take the form

$$\theta_i \ = \ \Sigma_{k=1}^n a_{ik}(q)dq^k + \Sigma_{k=1}^n b_{ik}(q)dv^k, \quad i = 1\cdots m. \qquad (5.3.20)$$

A constrained dynamical system is then a special (2nd order) vector field compatible with these constraints. The case of interest to us is when the θ_i are differentials of functions on TQ i.e., of the form $df(q,v)$ – in particular the case $f(q,v) = \sum_k a_{ik}(q)v^k$ so

$$\theta_i = \sum_k a_{ik}dv^k + \sum_{k,j}\ \frac{\partial a_{ik}}{\partial q^j}v^k dq^j\ . \qquad (5.3.21)$$

We now want to define reaction forces—the forces that keep the system on the constraint manifold and to show that in the "ideal" case such forces do no work.

5.3.2 Definition. *Given constraints θ_i as in (5.3.20) we define $P^*\theta_i(= \sum b_{ik} dq^k)$ to be the reaction forces.*

Note that in the case (5.3.21) this is just

$$P^*\theta_i = \sum_k a_{ik} dq^k. \tag{5.3.22}$$

5.3.3 Definition. *A set of constraints is said to be* ***admissible*** *if*

$$\mathrm{Dim\,span}\ P^*\theta_i = \mathrm{Dim\,span}\ \theta_i.$$

(This is equivalent to saying that the codistribution given by the θ_i has no horizontal covectors at any point, since the kernel of P^* is horizontal 1-forms. Another way to think of this: since horizontal vectors are linear combinations of vectors $\frac{\partial}{\partial q^i}$, the vectors $a_i = (a_{i1}, \cdots, a_{in})$ must be linear combination of the vectors $b_i = (b_{i1}, \cdots, b_{in})$ so that if one has some vectors in the span of the $\frac{\partial}{\partial q^i}$ there is also a component in the span of the $\frac{\partial}{\partial v^i}$.)

5.3.4 Definition. *A constraint is said to be* ***ideal*** *if it annihilates the fundamental vector field*

$$\Phi(= \sum_i v^i \frac{\partial}{\partial v^i} \text{locally}).$$

Then we have

5.3.5 Theorem. *If a set of constraints is admissible there exist special vector fields satisfying the constraints.*

Proof. Recall that the special vector fields are those satisfying $\tau X = \Phi$. Therefore we need to ask if the system

$$PX = \Phi \quad \theta_i(X) = 0, \quad i = 1, \cdots, m \tag{5.3.23}$$

is solvable.

Let

$$X = \sum_i v^i \frac{\partial}{\partial q^i} + f^i \frac{\partial}{\partial v^i}.$$

Since the constraints are admissible, by the argument above the vectors a^i are in the span of the vectors b^i. Now we require

$$\theta_i(X) = \sum_k b_{ik} f^k + \sum_k a_{ik} v^k = 0. \tag{5.3.24}$$

Thus we wish to know whether the system of linear equations

$$\sum_k b_{ik} f^k = -\sum_k a_{ik} v^k$$

in the f^i can be solved. But this is clear by the admissibility argument. In fact, since there are m conditions on the f^i there are at least $n - m$ independent special vector fields. ∎

A force is said to ***do no work*** if $\int F \cdot \; dq = 0$ along any curve in the configuration space.

5.3.6 Theorem. *If a constraint is ideal the corresponding constraint forces do no work (i.e., they are virtual forces.)*

Proof. Let ξ be a closed curve in Q, i.e., a map from S^1 to Q and let $\tilde{\xi}$ be its lift to TQ.

Then

$$\int_{\tilde{\xi}} \tau^* \theta_i = \int_{S^1} \left\langle \tau^* \theta_i, \dot{\tilde{\xi}} \right\rangle = \int_{S^1} \left\langle \theta_i, \tau \dot{\tilde{\xi}} \right\rangle$$

where $<,>$ is the natural pairing between forms and their duals.

But if $\xi = q(t)$ and $\tilde{\xi} = (q(t), \dot{q}(t))$ then

$$\dot{\tilde{\xi}} = \dot{q} \frac{\partial}{\partial q} + \ddot{q} \frac{\partial}{\partial v}. \tag{5.3.25}$$

Hence

$$\tau(\dot{\tilde{\xi}}) = \dot{q} \; \frac{\partial}{\partial v} = v \; \frac{\partial}{\partial v} = \Phi.$$

Thus we get

$$\int_{S^1} \langle \theta_i, \Phi \rangle \;=\; 0 \tag{5.3.26}$$

by ideality.

The usual arguments imply this integral is zero along any curve. ∎

Now we can show

5.3.7 Theorem. *Let $\theta_i, \; i = 1 \ldots m$ define a constraint distribution on TQ and let L be a nondegenerate Lagrangian with positive definite Hessian $\frac{\partial^2 L}{\partial v^i \partial v^j}$. Then there exists a special vector field X satisfying the Lagrange d'Alembert principle*

$$i_X \Omega_L = dH_L + \omega\,; \tag{5.3.27}$$

that is,

$$\Omega_L(X, Y) = (dH_L + \omega)(Y) \quad \forall \text{ vector field } Y,$$

where ω is the constraint force that ensures $\theta^i(X) = 0,\ i = 1, \ldots, m$. Locally the equations of motion are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial q^i}\right) - \frac{\partial L}{\partial q^i} = \sum \lambda_i \theta_i. \tag{5.3.28}$$

Proof. Since Ω_L is nondegenerate there is a map from 1-forms to vector fields π_L such that

$$\Omega_L\left(\pi_L(\rho), Y\right) = \rho(Y) \tag{5.3.29}$$

for Y an arbitrary tangent vector and ρ an arbitrary 1-form. Then

$$X = \pi_L(dH_L + \omega) \tag{5.3.30}$$

and we require

$$\Big\langle \theta_i, \pi_L(dH_L) + \pi_L(\omega)\Big\rangle = 0 \qquad i = 1, \ldots, m. \tag{5.3.31}$$

Now ω, by definition of the constraint force, needs to be a linear combination of $\tau^*\theta_i$, i.e., of the form $\sum_i \lambda_i a_{ik} dq_k$.That is, we have

$$\Big\langle \theta_i\,, \pi_L(dH_L)\Big\rangle = -\Big\langle \theta_i\,, \sum_i \lambda_i \pi_L \tau^* \theta_i \Big\rangle \quad i = 1, \ldots, m. \tag{5.3.32}$$

Since the constraints are admissible, τ^* preserves the dimension of the span of θ_i, but since π_L is positive definite the restriction of π_L to the span of the $\tau^*\theta_i$ is also positive definite and thus nondegenerate. Since the operator $\pi_L\tau^*$ is thus of full rank, we can solve the system for λ_i.

The local form of the equations follows from the general Lagrange theory above and of course agree with the coordinate form derived elsewhere. ■

5.5 The Momentum Equation.

The Nonholonomic Connection and Reconstruction

Here we discuss the application of the momentum equation to the problem of reconstructing paths on configuration space Q given a path in the base space Q/G. Many of the things already given in Chapter 5 will be given here from a slightly different point of view and with some additional details given.

In many systems the base space Q/G corresponds to the set of variables which are directly controlled by the application of control forces, and hence we can follow any path in Q/G by application of appropriate forces. It is therefore natural to focus on how these paths lift, as described by the constraints, the generalized momenta, and the momentum equation, to the full configuration space. The main tool tool to be discussed here is that of the *nonholonomic connection*, a synthesis of the mechanical and the kinematic connections.

The Unconstrained Case. We will begin by recalling the reconstruction procedure for *unconstrained* mechanical systems. As we discussed earlier for unconstrained mechanical systems with symmetries, the equations of motion are naturally described in using the principal bundle $Q \to Q/G$. In essence, the dynamical equations split into two pieces by using Hamilton's principle $\delta L = 0$ and dividing the variations into vertical variations and a set of complementary variations. The vertical variations lead to a set of conservation laws of the form

$$\frac{d}{dt}\langle \mathbb{F}L, \eta_Q \rangle = 0,$$

for all $\eta \in \mathfrak{g}$. These equations are equivalent to the Euler-Poincaré equations when the Euler-Lagrange equations are written in a local trivialization. As we mentioned above, the mechanical connection is related to the momentum map and the locked inertia tensor by

$$\mathcal{A}(q) \cdot v_q = \mathbb{I}^{-1}(q) J(v_q).$$

Given a path in the base space Q/G, we can now use the connection to reconstruct the motion of the system in the full space Q. The conservation law can be written as

$$\mathcal{A}(q) \cdot \dot{q} = \mathbb{I}^{-1}(q) J(\dot{q}) = \mathbb{I}^{-1}(q)\mu$$

where $\mu \in \mathfrak{g}^*$ is a (constant) momentum. If we choose a local trivialization of the bundle with coordinates $q = (r, g) \in (Q/G) \times G$ (locally), the conservation law becomes

$$\mathcal{A}(q) \cdot \dot{q} = \mathrm{Ad}_g(g^{-1}\dot{g} + \mathcal{A}_{\mathrm{loc}}(r)\dot{r}) = (\mathrm{Ad}_g \mathbb{I}_{\mathrm{loc}}^{-1}(r) \mathrm{Ad}_g^*) \cdot \mu$$

where $\mathbb{I}_{\mathrm{loc}}$ is the local expression for the locked inertia tensor written as a function over Q/G. Rearranging this equation, we see that the group variables evolve according to

$$\dot{g} = g\left(-\mathcal{A}_{\mathrm{loc}}(r)\dot{r} + \Omega\right) \tag{5.5.16}$$

where $\Omega = \mathbb{I}_{\mathrm{loc}}^{-1}(r)p$ is the body angular velocity and where $p = \mathrm{Ad}_g^* \cdot \mu$ is the body angular momentum. Note that the variables p (or Ω if one is doing the Lagrangian point of view) are to be included amongst the variables in the reduced phase space. Thus, given a path $r(t)$ in the base variables, a motion in the body angular momentum (p) space or velocity (Ω) space, and an initial condition for the group variables, we can reconstruct the motion in the group and hence on the entire space, as in Marsden, Montgomery, and Ratiu [1990]. Finally, we reiterate a basic fact from this discussion: *the body angular velocity* $\Omega = \xi + \mathcal{A}_{\rm loc}(r)\dot{r}$ *(where* $\xi = g^{-1}\dot{g}$*) is the local representative of the vertical part of the velocity vector* $\dot{q}$*.*

If nonholonomic constraints are present, it is still possible to reconstruct the path in the group variables given the path in the base. This is useful in control applications since it allows us to study the motion of the system without considering the full equations of motion. We break the following discussion into three cases: purely kinematic constraints, horizontal symmetries, and the general case. The purely kinematic case occurs when the constraint distribution complements the symmetry group orbit. In this case, it is clear that we do not get any nontrivial components to the momentum equation and that the constraint distribution itself defines a principal connection.

The Principal or Purely Kinematic Case. Recall that a set of nonholonomic constraints is said to be *purely kinematic* if the constraints define a connection on a principal bundle and that this situation occurs when the constraint distribution is G-invariant and the tangent space to the group orbit forms a complement to the constraint distribution; that is, the subbundle with the fibers $\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \{0\}$ for all $q \in Q$. What this really means is that there are no momentum equations in this case and that correspondingly there is no analogue of the body angular momentum or velocity, as there was in the preceding discussion of unconstrained systems. In particular, relative to a local trivialization $q = (r, g)$ the constraints can be written as

$$A(q)\dot{q} = \left[\mathrm{Ad}_g(g^{-1}\dot{g} + \mathcal{A}_{\rm loc}(r)\dot{r})\right]_Q = 0.$$

The motion in the fibers is thus given by

$$\dot{g} = -g\mathcal{A}_{\rm loc}(r)\dot{r}$$

and we can reconstruct the group motion given the trajectory in Q/G. In this case, as we saw previously, the equations reduce to second order equations for r; that is, to second order equations on Q/G. The motion on the full space is then determined by the solution to these reduced equations followed by first order equations for the group variables.

This can be said a slightly different way: in the case of purely kinematic constraints, the kinematic connection replaces the mechanical connection to determine the motion in the fibers. This situation occurs only when the constraint distribution $\mathcal{D}$ and the vertical subspace $T_q(\mathrm{Orb}(q))$ are such that $T_qQ = T_q(\mathrm{Orb}(q)) \oplus \mathcal{D}_q$, so that $\mathcal{D}_q$ can be taken as the horizontal space for a connection. Thus the conservation law which would govern the motion in the group variables if no constraints were present is replaced by the motion dictated by the constraints. See Koiller [1992] for a further discussion of the purely kinematic case, including a description of reduction in that context. This reduction result can be obtained as a special case of the results of Cendra, Marsden, and Ratiu [2001b], where it is shown how to reduce the horizontal part of the variational principal relative to any connection.

The Case of Horizontal Symmetries. A second case in which it is possible to lift the motion from the base Q/G to the fibers using a connection is when there are enough horizontal symmetries such that they and the constraints interact in a complementary fashion. This situation occurs, for example, when there is a subgroup of G whose action on the configuration space satisfies the constraints. We call a symmetry of this type a *horizontal symmetry* (relative to the kinematic constraints). When horizontal symmetries are present, the motion in the group variables can be reconstructed by combining the kinematic constraints with the conservation laws corresponding to the horizontal symmetries. This is the case for the ball on the rotating plate.

We begin by restricting ourselves to the case when $\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ$ for all $q \in Q$ and we assume that there exists a a subgroup $H \subset G$ such that $\xi_Q \in \mathcal{D}$ for all $\xi \in \mathfrak{h}$ and $\mathcal{D}_q \cap T_q(\mathrm{Orb}_G(q)) = T_q(\mathrm{Orb}_H(q))$. We call H the group of horizontal symmetries and define the momentum map $J_H : TQ \to \mathfrak{h}^* \subset \mathfrak{g}^*$ as

$$\langle J_H(v_q), \xi \rangle = \langle \mathbb{F}L(v_q), \xi_Q \rangle \qquad \xi \in \mathfrak{h}.$$

For a Lagrangian of the form kinetic energy minus potential energy we can write the generalized momenta as linear functions of the velocity and these generalized momenta are constant along solution curves since $\xi^q = \xi \in \mathfrak{h}$ is constant. Thus we have

$$\langle J_H(q) \cdot \dot{q}, \xi \rangle = \langle \mu, \xi \rangle \qquad \xi \in \mathfrak{h}$$

where $\mu \in \mathfrak{h}^*$ is a constant and we see that the generalized momentum has the form of an affine constraint

$$J_H(q) \cdot \dot{q} = \mu. \tag{5.5.17}$$

To reconstruct the motion in the fibers, we build a connection on $Q \to Q/G$ by augmenting the kinematic constraints with the conservation law. Let $\mathbb{I}(q) : \mathfrak{h} \to \mathfrak{h}^*$ be the locked inertia tensor relative to $\mathfrak{h}$, defined by

$$\langle \mathbb{I}(q)\xi, \eta \rangle = \langle\langle \xi_Q, \eta_Q \rangle\rangle \qquad \xi, \eta \in \mathfrak{h}.$$

We define a map $A^{\mathrm{sym}} : TQ \to \mathcal{S}$ as

$$A^{\mathrm{sym}}(v_q) = \left(\mathbb{I}^{-1}(q) J_H(v_q)\right)_Q \tag{5.5.18}$$

and the conservation law (5.5.17) can be rewritten as an affine constraint

$$A^{\mathrm{sym}}(\dot{q}) = \left(\mathbb{I}^{-1}(q)\mu\right)_Q . \tag{5.5.19}$$

The one-form A^{sym} takes values in $\mathcal{S}_q = T_q(\mathrm{Orb}_H(q))$ and is equivariant with respect to the full group action since the kinetic energy metric is invariant and the momentum map is equivariant. It also follows from the definition of the momentum map that A^{sym} is a projection onto $\mathcal{S}$ and hence it maps vectors on $\mathcal{S}_q$ to themselves.

By assumption, the constraint distribution $\mathcal{D}$ is invariant and if we choose a subspace $\mathcal{U}_q \subset T_q(\mathrm{Orb}(q))$ such that $T_q(\mathrm{Orb}(q)) = \mathcal{U}_q \oplus \mathcal{S}_q$ then we can represent the constraints using a $\mathcal{U}$ valued one-form $A^{\mathrm{kin}} : TQ \to \mathcal{U}_q$ where A^{kin} satisfies the following conditions:

$$\left.\begin{array}{c} A^{\mathrm{kin}}(v_q) = 0 \quad \text{if and only if} \quad v_q \in \mathcal{D}_q \\ A^{\mathrm{kin}}(v_q) = v_q \quad \text{for all} \quad v_q \in \mathcal{U}_q \\ \Phi_{g*} A^{\mathrm{kin}} = A^{\mathrm{kin}} \Phi_{g*}. \end{array}\right\} \tag{5.5.20}$$

We now combine the two mappings to form a new mapping $A : TQ \to T\,\mathrm{Orb}$, where $T\,\mathrm{Orb}$ denotes the union of the tangent spaces to the group orbits, that is, to the vertical bundle for the projection $Q \to Q/G$, as follows:

$$A = A^{\mathrm{kin}} + A^{\mathrm{sym}}. \tag{5.5.21}$$

The mapping $A : TQ \to T\,\mathrm{Orb}$ is an equivariant Ehresmann connection on the bundle $Q \to Q/G$ and hence we can write

$$A(v_q) = (\mathcal{A}(v_q))_Q,$$

where $\mathcal{A} : TQ \to \mathfrak{g}$ is a principal connection. To see that A is an Ehresmann connection it suffices to show that it is a projection on $\mathcal{U}_q$ and $\mathcal{S}_q$. This follows immediately from the fact that A^{sym} and A^{kin} are equivariant projections onto $\mathcal{S}$ and $\mathcal{U}$ respectively and $T_q(\mathrm{Orb}(q)) = \mathcal{S}_q \oplus \mathcal{U}_q$. Equivariance follows directly from the equivariance of $\mathcal{U}_q$ and $\mathcal{S}_q$ and the existence of $\mathcal{A}$ follows from general properties of equivariant Ehresmann connections.

5.5.9 Definition. *We call the map $A : TQ \to T_q(\mathrm{Orb}(q))$ defined by equations* (5.5.19)–(5.5.21) *the* ***nonholonomic connection*** *(in the case of horizontal symmetries).*

Notice that the nonholonomic connection in the case of horizontal symmetries reduces to the kinematic connection in the purely kinematic case and the mechanical connection in the unconstrained case. See Figure 5.5.1

The overall motion of the system satisfies

$$A(q) \cdot \dot{q} = (\mathbb{I}^{-1}(q)\mu)_Q \tag{5.5.22}$$

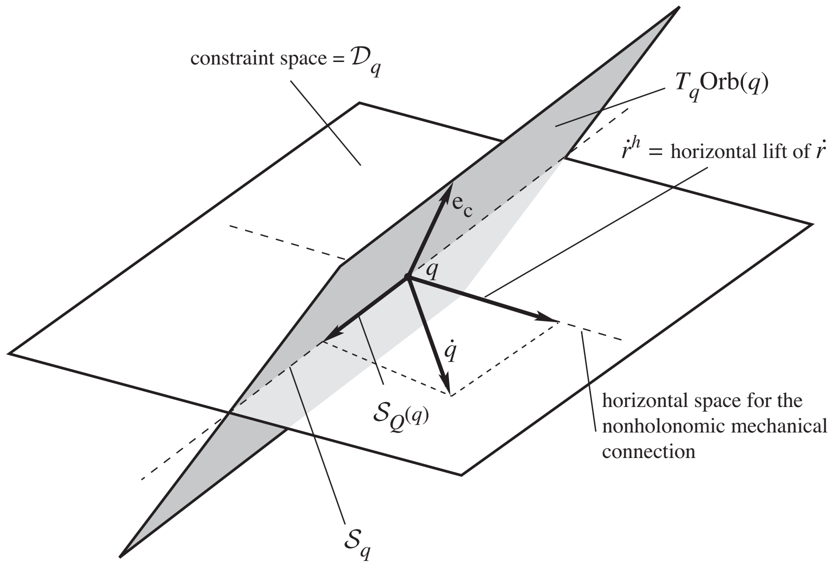

FIGURE 5.5.1. The decomposition of $\dot{q}$ into vertical and horizontal pieces relative to the nonholonomic connection.

which has the form of an affine constraint. The locked inertia tensor relative to $\mathfrak{h}$ satisfies

$$\mathbb{I}(g \cdot q) = \mathrm{Ad}^*_{g^{-1}}\mathbb{I}(q)\mathrm{Ad}_{g^{-1}}$$

and hence in general the affine part of the constraint (5.5.22) is *not* equivariant since

$$\mathbb{I}^{-1}(g \cdot q)\mu = \mathrm{Ad}_g\mathbb{I}^{-1}(q)\mathrm{Ad}^*_g\mu \neq \mathrm{Ad}_g(\mathbb{I}^{-1}(q)\mu).$$

This lack of invariance of the affine portion, as in the unconstrained case, would cause problems in the construction of a principal connection if one tried to make full use of the conservation laws by holding μ fixed. On the other hand, the actual reduced variables correspond to the body angular velocity or momentum, and in these variables, equivariance is restored. Let us be more specific: Equation (5.5.22) describes how to lift paths from the base space Q/G to the full space Q. This is most easily seen relative to a local trivialization $q = (r, g)$, where the constraints can be written as

$$\mathcal{A}(q) \cdot \dot{q} = \mathrm{Ad}_g(g^{-1}\dot{g} + \mathcal{A}_{\mathrm{loc}}(r)\dot{r}) = \mathrm{Ad}_g\mathbb{I}^{-1}_{\mathrm{loc}}(r)\mathrm{Ad}^*_g\mu$$

where $\mathrm{Ad}_g\mathbb{I}^{-1}_{\mathrm{loc}}(r)\mathrm{Ad}^*_g\mu$ is the $\mathfrak{g}^{\mathcal{D}}$-valued function associated with the constant momentum $\mu \in \mathfrak{h}^*$. This equation can be rewritten as

$$\dot{g} = g(-\mathcal{A}_{\mathrm{loc}}(r)\dot{r} + \mathbb{I}^{-1}_{\mathrm{loc}}(r)\mathrm{Ad}^*_g\mu)$$

which shows how the path $r(t) \in Q/G$ lifts to the fibers.

Noting that $\mathrm{Ad}_g^*\mu = p$ is the body angular momentum and $\mathbb{I}_{\mathrm{loc}}^{-1}(r)p = \Omega$ is the corresponding body angular velocity, which may be regarded as a dynamical variable in its own right, then the reconstruction equation takes the form

$$\dot{g} = g(-\mathcal{A}_{\mathrm{loc}}(r)\dot{r} + \Omega). \tag{5.5.23}$$

This equation again has the form $\dot{g} = g\xi$ and where $\xi = -\mathcal{A}_{\mathrm{loc}}(r)\dot{r} + \Omega$ has been determined by equations of motion that themselves are independent of the group variable. This form, rather than the form in which the momentum has been set equal to a constant shows the decoupling from the group variables most clearly. As we saw before, and will do more generally below, it is the variable ξ rather than the body angular velocity variable that evolves by means of a component of the Euler-Poincaré equation. On the other hand, it is Ω that is the vertical variable relative to the nonholonomic connection

$$\mathrm{ver}_q\dot{q} = \Omega = \mathcal{A}_{\mathrm{loc}}(r)\dot{r} + g^{-1}\dot{g},$$

which is an instance of the general coordinate expression for the vertical part of a principal connection. As we shall see in a moment, this point of view generalizes to the case of nonhorizontal symmetries.

The preceding equations only hold when $\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ$ and $\mathcal{D}_q \cap T_q(\mathrm{Orb}_G(q)) = T_q(\mathrm{Orb}_H(q))$. If we drop the second restriction, then the reconstruction procedure must be modified to account for the interaction between the constraints and the symmetries. The developments below will include this more general situation.

Finally we end with a notational remark. In the general nonholonomic case, as we have seen, the momentum map need not be conserved. In any case, even if it is, the momentum in body representation, p is not constant.

The Nonholonomic Connection

We now consider the most general case, where the symmetries are not necessarily horizontal. Although it is not needed for everything that we will be doing, the examples and the theory are somewhat simplified if we make the following assumption:

Dimension Assumption. *The constraints and the orbit directions span the entire tangent space to the configuration space:*

$$\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ \tag{5.5.24}$$

If this condition holds, we shall say that we are in the ***principal case****.*

In this case, the momentum equation can be used to augment the constraints and provide a connection on $Q \to Q/G$. Let $J^{\mathrm{nhc}} : TQ \to (\mathfrak{g}^{\mathcal{D}})^*$

be the nonholonomic momentum map,

$$\langle J^{\mathrm{nhc}}(q) \cdot \dot{q}, \xi^q \rangle = \langle \mathbb{F}L, \xi^q_Q \rangle$$

and define, as before, a map $A^{\mathrm{sym}}_q : T_qQ \to \mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q))$ given by

$$A^{\mathrm{sym}}_q(v_q) = (\mathbb{I}^{-1} J^{\mathrm{nhc}}(v_q))_Q \tag{5.5.25}$$

This map is equivariant and a projection onto $\mathcal{S}_q$. Here $\mathbb{I} : \mathfrak{g}^{\mathcal{D}} \to (\mathfrak{g}^{\mathcal{D}})^*$ is the locked inertia tensor relative to $\mathfrak{g}^{\mathcal{D}}$; it is defined in the same way as before.

If we now choose $\mathcal{U}_q \subset T_q(\mathrm{Orb}(q))$ such that $T_q(\mathrm{Orb}(q)) = \mathcal{S}_q \oplus \mathcal{U}_q$ then we can synthesize a connection which encodes both the constraints and the momenta, as before. Let $A^{\mathrm{kin}}_q : T_qQ \to \mathcal{U}_q$ be a $\mathcal{U}_q$ valued form that projects $\mathcal{U}_q$ onto itself and maps $\mathcal{D}_q$ to zero; for example, it can be given by orthogonal projection relative to the kinetic energy metric (this will be our default choice). The constraints plus momentum equation can thus be written as

$$\begin{aligned} A^{\mathrm{kin}}(q) \cdot \dot{q} &= 0 && \text{(constraints)} \\ A^{\mathrm{sym}}(q) \cdot \dot{q} &= (\mathbb{I}^{-1}(q)p)_Q && \text{(momenta),} \end{aligned}$$

where $p \in (\mathfrak{g}^{\mathcal{D}})^*$ is the time dependent momentum defined by

$$p = \left\langle J^{\mathrm{nhc}}(q) \cdot \dot{q}, \xi^q \right\rangle .$$

5.5.10 Definition. *Under the dimension assumption in equation (5.5.24), and the assumption that the Lagrangian is of the form kinetic minus potential energies, the* ***nonholonomic connection*** $\mathcal{A}$ *is the connection on the principal bundle* $Q \to Q/G$ *whose horizontal space at the point* $q \in Q$ *is given by the orthogonal complement to the space* $\mathcal{S}_q$ *within the space* $\mathcal{D}_q$.

Under the assumption that the distribution is invariant (condition (S1)), and from the fact that the group action preserves orthogonality, it follows that the distribution $\mathcal{S}$ and the horizontal spaces transform to themselves under the group action. Thus, we get:

5.5.11 Proposition. *Under the assumptions in the previous definition and the condition (S1), the nonholonomic connection is a principal connection.*

Using the preceding expressions, an expression for the nonholonomic connection as an Ehresmann connection (and hence also as a principal connection) is given by our earlier calculations. In fact, one can readily check that the following proposition holds:

5.5.12 Proposition. *The nonholonomic connection regarded as an Ehresmann connection is given by*

$$A = A^{\mathrm{kin}} + A^{\mathrm{sym}}. \tag{5.5.26}$$

When the connection is regarded as a principal connection (i.e., takes values in the Lie algebra rather than the vertical space) we will use the symbol $\mathcal{A}$.

The nonholonomic connection defined here agrees with the definition in the horizontal case. (In making this comparison, note that in the general definition of the connection, we do *not* fix the value of μ but rather let it be determined by the point v_q at which the connection is evaluated.)

The affine constraint $A(q) \cdot \dot{q} = (\mathbb{I}^{-1}(q) \cdot p)_Q$ describes the lifting of paths from the base. The formula for the nonholonomic connection is given in terms of A^{kin}, which depends on the choice of complement $\mathcal{U}_q$ to $\mathcal{S}_q$ within the tangent space to the orbit. However, it is easily seen that $A : TQ \to T\,\mathrm{Orb}$ is independent of this choice, as it must be since the definition of the nonholonomic connection was manifestly independent of this choice.

Special Cases. Various special cases can be conveniently classified by the generic and extreme ways the subspaces in the preceding figure intersect. For example, the purely kinematic case is when the space $\mathcal{S}_q$ is zero dimensional. The extreme case in which the tangent space to the orbit is a subset of the space of constraints is itself an extreme case of that of horizontal symmetries, etc. These different cases we have discussed are summarized in Table 5.5.1.

Case	Conditions	Connection
Unconstrained	$\mathcal{D}_q = T_qQ$	$\mathcal{A}^{\mathrm{sym}}(\dot{q}) = \mathbb{I}^{-1}J(\dot{q})$
Purely kinematic	$\mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \{0\}$	$\mathcal{A}^{\mathrm{kin}}(\dot{q}) = 0$
Horizontal symmetries	$\mathcal{D}_q \cap T_q(\mathrm{Orb}_G(q))$ $= T_q(\mathrm{Orb}_H(q))$	$\mathcal{A}^{\mathrm{sym}}(\dot{q}) + \mathcal{A}^{\mathrm{kin}}(\dot{q})$ $= \mathbb{I}^{-1}J_H(\dot{q})$
General principal bundle case	$\mathcal{D}_q + T_q(\mathrm{Orb}(q)) = T_qQ$	$\mathcal{A}^{\mathrm{sym}}(\dot{q}) + \mathcal{A}^{\mathrm{kin}}(\dot{q})$ $= \mathbb{I}^{-1}J^{\mathrm{nhc}}(\dot{q})$

TABLE 5.5.1. Special cases of the nonholonomic connection (principal case).

In addition to these possibilities, one can also consider the case where $\mathcal{D}_q + T_q(\mathrm{Orb}(q)) \neq T_qQ$. When this happens the base space for the Ehresmann connection can no longer be chosen as Q/G and hence a bigger base space must be chosen. However, the basic constructions still hold with the momentum augmenting the constraints to give a synthesized connection.

Within this overall framework, reduction is also possible in certain cases. For example, in the purely kinematic case, Koiller [1992] showed that the dynamics of the system drop to the base space Q/G. Similarly, in the case of horizontal symmetries, we have discussed the situation above. The general case will be discussed below and the reduced equations computed. In the general case, the reduced equations will define a dynamical system on the space $\mathcal{D}/G$, and the reconstruction problem, which we have largely discussed already, will be the problem of lifting the dynamics from $\mathcal{D}/G$ back to the space $\mathcal{D} \subset TQ$.

5.8 Computation of the Reduced Lagrange d'Alembert Equations

Here we compute explicitly the nonholonomic equations on the reduced space $\mathcal{D}/G$. The strategy is to explore the equations of motion, split according to the nonholonomic connection that was constructed above. Throughout this section we make the dimension assumption of §5.7 of the main text, so that the nonholonomic connection is a principal connection. Without this assumption, one would have to assume an additional bundle structure. We avoid this for simplicity and because the dimension assumption holds in all our examples and other related ones we know about (such as the bicycle, the rolling ellipsoid, etc.).

The Momentum Equation in an Orthogonal Body Frame. We shall first compute the reduced form of the momentum equation that will be one of the sets of equations comprising the reduced Lagrange d'Alembert equations. This splitting of the equations is associated with breaking up the variations that go into the Lagrange d'Alembert principle into vertical and horizontal parts relative to the nonholonomic connection. To do this, we make the assumption, that the initial Lagrangian is of the form of kinetic minus potential energy; in particular, the metric structure defined by the kinetic energy will be used. Using the kinetic energy metric, we choose our moving basis $e_c(q)$ to be orthogonal; that is, the corresponding generators $[e_c(q)]_Q$ are orthogonal in the given kinetic energy metric. (Actually, all that is needed is that the vectors in the set of basis vectors corresponding to the subspace $\mathcal{S}_q$ be orthogonal to the remaining basis vectors.) The metric tensor will be denoted by g_{ij}.

We begin by recalling the decompositions defined by the nonholonomic connection described earlier. Given a velocity vector $\dot{q}$ that satisfies the constraints, we orthogonally decompose it into a piece in $\mathcal{S}_q$ and an orthogonal piece denoted $\dot{r}^h$. We regard $\dot{r}^h$ as the horizontal lift of a velocity vector $\dot{r}$ on shape space; recall that in a local trivialization, the horizontal lift to the point (r, g) is given by

$$\dot{r}^h = (\dot{r}, -\mathcal{A}_{\text{loc}}\dot{r}) = (\dot{r}^\alpha, -\mathcal{A}^a_\alpha \dot{r}^\alpha)$$

where $\mathcal{A}^a_\alpha$ are the components of the nonholonomic connection (recall that it is a principal connection) in a local trivialization.

We will denote the decomposition of $\dot{q}$ by

$$\dot{q} = \Omega_Q(q) + \dot{r}^h,$$

so that for each point q, ω is an element of the Lie algebra and represents the spatial angular velocity of the locked system. Note that in this expression, the constraints are implicitly included. In a local trivialization, we can write, at a point (r, g)

$$\Omega = \mathrm{Ad}_g(\Omega_{\mathrm{loc}})$$

so that Ω_{loc} represents the body angular velocity. Thus,

$$\Omega_{\mathrm{loc}} = \mathcal{A}_{\mathrm{loc}}\dot{r} + \xi$$

and, at each point q, the constraints are that Ω belongs to $\mathfrak{g}^q$, *i.e.,*

$$\Omega \in \mathrm{span}\{e_1(r), e_2(r), \ldots, e_m(r)\}.$$

As noted above, the vector $\dot{r}^h$ need not be orthogonal to the whole orbit, just to the piece $\mathcal{S}_q$. Even if $\dot{q}$ does not satisfy the constraints we can decompose it into three parts according to the figure and write

$$\dot{q} = \Omega_Q(q) + \dot{r}^h = \Omega_Q^{\mathrm{nh}}(q) + \Omega_Q^{\perp}(q) + \dot{r}^h,$$

where Ω_Q^{nh} lies in the space $\mathcal{S}_q$, that is, it satisfies the constraints, and is perpendicular within $T_q\,\mathrm{Orb}$ to $\Omega_Q^{\perp}$. The relation $\Omega_{\mathrm{loc}} = \mathcal{A}_{\mathrm{loc}}\dot{r} + \xi$ is valid even if the constraints do not hold; also note that this decomposition of Ω corresponds to the decomposition of the nonholonomic connection given by $A = A^{\mathrm{kin}} + A^{\mathrm{sym}}$.

We begin with the momentum equation in body representation, which we recall here for convenience:

$$\frac{d}{dt}p_b = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_b] + \frac{\partial e_b}{\partial r^\alpha}\dot{r}^\alpha \right\rangle. \tag{5.8.1}$$

This equation is one of the reduced equations since it manifestly decouples from the group variables. We shall now work out this equation in coordinates.

As above we make the following index and summation conventions

1. The first batch of indices range from 1 to m corresponding to the symmetry directions along constraint space. These indices will be denoted $a, b, c, d, \ldots$ and a summation from 1 to m will be understood.

2. The second batch of indices range from $m+1$ to k corresponding to the symmetry directions not aligned with the constraints. Indices for this range or for the whole range 1 to k will be denoted by $a', b', c', \ldots$ and the summations will be given explicitly.

3. The indices $\alpha, \beta, \ldots$ on the shape variables r range from 1 to σ. Thus, σ is the dimension of the shape space Q/G and so $\sigma = n - k$. The summation convention for these indices will be understood.

We shall need the following calculation:

5.8.6 Proposition. *In a local trivialization we have*

$$\left\langle \frac{\partial l}{\partial \xi}, \eta \right\rangle = I_{ac}(r)\Omega^a \eta^c + \sum_{a'=m+1}^{k} \lambda_{a'\alpha} \eta^{a'} \dot{r}^\alpha = p_c \eta^c + \sum_{a'=m+1}^{k} \lambda_{a'\alpha} \eta^{a'} \dot{r}^\alpha. \tag{5.8.2}$$

In this equation, the partial derivatives of l are evaluated at a point $(r, \dot{r}, \xi)$ satisfying the constraints (that is, the corresponding $\Omega_{\mathrm{loc}} = \xi + \mathcal{A}_{\mathrm{loc}} \dot{r}$ lies in $\mathfrak{g}^q$) and η is an arbitrary element of $\mathfrak{g}$. Also,

$$p_b = I_{ab}(r)\Omega^a,$$

where $I_{ab}(r)$ are the coefficients of the locked inertia tensor $\mathbb{I}_{\mathrm{loc}}(r)$ in a local trivialization (recall from the last section that the locked inertia tensor has indices that range only over the first batch), and where

$$\lambda_{a'\alpha} = l_{a'\alpha} - \sum_{b'=1}^{k} l_{a'b'} \mathcal{A}_\alpha^{b'} := \frac{\partial l}{\partial \xi^{a'} \partial \dot{r}^\alpha} - \sum_{b'=1}^{k} \frac{\partial l}{\partial \xi^{a'} \partial \xi^{b'}} \mathcal{A}_\alpha^{b'}, \tag{5.8.3}$$

for $a' = m+1, \ldots k$.

Proof. We denote the kinetic energy metric on T_qQ by $\langle\!\langle\, , \rangle\!\rangle_q$. The corresponding metric on $\mathfrak{g}$ restricted to the subspace $\mathfrak{g}^q$ gives the locked inertia tensor as we saw before.

The kinetic energy is given as follows, without the assumption that $\dot{q}$ satisfies the constraints:

$$\begin{aligned} K(q, \dot{q}) &= \frac{1}{2} \left\langle\!\left\langle \Omega_Q^{\mathrm{nh}} + \Omega_Q^\perp + \dot{r}^h, \Omega_Q^{\mathrm{nh}} + \Omega_Q^\perp + \dot{r}^h \right\rangle\!\right\rangle_q \\ &= \frac{1}{2} \left\langle\!\left\langle \Omega_Q^{\mathrm{nh}}, \Omega_Q^{\mathrm{nh}} \right\rangle\!\right\rangle_q + \left\langle\!\left\langle \Omega_Q^\perp, \dot{r}^h \right\rangle\!\right\rangle_q \\ &\quad + \frac{1}{2} \left\langle\!\left\langle \Omega_Q^\perp, \Omega_Q^\perp \right\rangle\!\right\rangle_q + \frac{1}{2} \left\langle\!\left\langle \dot{r}^h, \dot{r}^h \right\rangle\!\right\rangle_q , \end{aligned} \tag{5.8.4}$$

where we have suppressed the q dependence of $\Omega_Q(q)$ for simplicity.

Now we pass to a local trivialization and remove the explicit g dependence. We change variables to $(r, \dot{r}, \Omega)$ by the transformation $\Omega = \xi + \mathcal{A}_{\mathrm{loc}} \dot{r}$, which is valid even if the constraints are not satisfied. The partial derivatives with respect to Ω equal those with respect to ξ (evaluated at the corresponding points).

To form the reduced Lagrangian, we substitute $\dot{r}^h = (\dot{r}^\alpha, -\mathcal{A}^a_\alpha \dot{r}^\alpha)$ into the second term and arrive at

$$\frac{1}{2} I_{ac} \Omega^a \Omega^c + \sum_{a'=m+1}^{k} l_{a',\alpha} \Omega^{a'} \dot{r}^\alpha - \sum_{a',c'=m+1}^{k} l_{a'c'} \Omega^{a'} \mathcal{A}^{c'}_\alpha \dot{r}^\alpha + \chi$$

where Ω^a and $\Omega^{a'}$ are the components of Ω^{nh} and $\Omega^\perp$ respectively, where the subscripts on the l denote the corresponding partial derivatives, as above, and where χ corresponds to the last two terms in (5.8.4), which will vanish when the partials are taken with respect to Ω at $\Omega^\perp = 0$. It should now be clear that the derivatives of this expression evaluated at $\Omega^\perp = 0$ are as stated in the proposition. ■

The coefficients $\lambda_{a'\alpha}$ measure the failure of the horizontal space for the nonholonomic connection to be orthogonal to the tangent space to the orbit.

Next, for each b such that $1 \le b \le m$, we write out the components of the remaining expression in (5.8.1):

$$\begin{aligned}
[\xi, e_b] + \frac{\partial e_b}{\partial r^\alpha} \dot{r}^\alpha &= [\Omega - \mathcal{A}_{\mathrm{loc}} \dot{r}, e_b] + \frac{\partial e_b}{\partial r^\alpha} \dot{r}^\alpha \\
&= \sum_{c'=1}^{k} C^{c'}_{ab} \Omega^a e_{c'} - \sum_{a',c'=1}^{k} C^{c'}_{a'b} \mathcal{A}^{a'}_\alpha \dot{r}^\alpha e_{c'} \\
&\quad + \sum_{c'=1}^{k} \gamma^{c'}_{b\alpha} \dot{r}^\alpha e_{c'}
\end{aligned} \tag{5.8.5}$$

where the symbols such as $C^{c'}_{a'b}$ are the corresponding components of the structure constants in the given basis and where we have written

$$\frac{\partial e_b}{\partial r^\alpha} = \sum_{c'=1}^{k} \gamma^{c'}_{b\alpha} e_{c'} \tag{5.8.6}$$

and

$$\mathcal{A}_{\mathrm{loc}} \dot{r} = \sum_{a'=1}^{k} \mathcal{A}^{a'}_\alpha \dot{r}^\alpha e_{a'}. \tag{5.8.7}$$

Substituting (5.8.2), (5.8.5), (5.8.6) and (5.8.7) into (5.8.1), we arrive at the following.

5.8.7 Proposition. *The momentum equation in an orthogonal body frame is given as follows:*

$$\frac{d}{dt} p_b = C^c_{ab} I^{ad} p_c p_d + \mathcal{D}^c_{b\alpha} \dot{r}^\alpha p_c + \mathcal{D}_{\alpha\beta b} \dot{r}^\alpha \dot{r}^\beta, \tag{5.8.8}$$

where

$$\mathcal{D}^c_{b\alpha} = \sum_{a'=1}^{k} -C^c_{a'b}\mathcal{A}^{a'}_\alpha + \gamma^c_{b\alpha} + \sum_{a'=m+1}^{k} \lambda_{a'\alpha} C^{a'}_{ab} I^{ac}, \tag{5.8.9}$$

$$\mathcal{D}_{\alpha\beta b} = \sum_{a'=m+1}^{k} \lambda_{a'\alpha} \left(-C^{a'}_{ab}\mathcal{A}^a_\beta + \gamma^{a'}_{b\beta} \right). \tag{5.8.10}$$

In the case of the snakeboard, the subspace $\mathfrak{g}^q$ is one dimensional as we shall see, and the following corollary applies.

5.8.8 Corollary. *If the subspace $\mathfrak{g}^q$ is either one dimensional or abelian, then the first term on the right hand side of* (5.8.8)*, which is quadratic in p, is zero.*

Another notable special case is the following, which will be used in the example of a constrained particle in $\mathbb{R}^3$ to produce a nontrivial parallel transport equation.

5.8.9 Corollary. *If $\mathfrak{g}$ is abelian, and if the horizontal space is (kinetic energy metric) orthogonal to the group orbit, then the momentum equation is in the form of a parallel transport equation over the curve $r(t)$ in shape space:*

$$\frac{d}{dt} p_b = \gamma^c_{b\alpha} \dot{r}^\alpha p_c.$$

We observe that the parallel transport form of the equations is characterized by the vanishing of the terms in the momentum equation that are purely quadratic in $\dot{r}$ and in p. This situation is important in understanding the complete integrability of some systems, such as Routh's problem of the rolling ball in a surface of revolution; cf Zenkov [1995].

The Reduced Equations. We now are in a position to put several parts of the preceding discussions together. As we saw above, the momentum equation in body representation decouples from the group variables themselves, which is important for the reconstruction strategy. On the other hand, this is a local representation for the intrinsic equations on the space $\mathcal{D}/G$. As we mentioned before, it is convenient to write them in local representation in terms of the variables Ω and $\dot{r}$ for several reasons:

1. This split of the equations corresponds to a global intrinsic split of the Lagrange–d'Alembert principle according to the nonholonomic connection (we emphasize that there is some freedom here; other connections can be used in its place).

2. This split enables us to use the (locked) body angular velocity Ω as a basic variable instead of ξ since it has better diagonalization properties for the kinetic energy and will ultimately be more useful

for purposes of stability analyses; these two variables are related by the velocity shift given by the nonholonomic connection:

$$\Omega_{\rm loc} = \mathcal{A}_{\rm loc}\dot{r} + \xi.$$

We will show that the equations of motion can be written (using a local trivialization) as three systems of equations, namely

- The constraint equations
- The reduced Euler-Lagrange equations using the nonholonomic connection for the variable $\dot{r}$
- The momentum equation (of Euler-Poincaré type) in body representation.

We formulate the reduced Lagrange–d'Alembert equations under the assumptions of Proposition 5.5.11. In this context, the Lagrange–d'Alembert principle may be broken up into two principles by breaking the variations δq into two parts, namely parts that are horizontal with respect to the nonholonomic connection and parts that are vertical (but still in $\mathcal{D}$). We will use as variables, $(r^\alpha, \dot{r}^\alpha, \Omega^a)$ where $(r, \dot{r})$ are variables in the base and where Ω is the vertical part (the locked body angular velocity). Let $l_c(r, \dot{r}, \Omega)$ denote the reduced Lagrangian written in terms of these variables as before; the subscript c is used to indicate the fact that Ω is confined to the constraint subspace $\mathfrak{g}^q$. Use the *orthogonal* basis $e_1(r), e_2(r), \ldots, e_m(r), e_{m+1}(r), \ldots e_k(r)$ introduced for the momentum equation in body representation (recall that this means that the first m elements are orthogonal as are the second $k-m$ elements but that the two sets need not be orthogonal to each other). Let

$$p_b(r, \dot{r}, \Omega) = \left\langle \frac{\partial l_c}{\partial \Omega}, e_b(r) \right\rangle, \quad b = 1, \ldots, m.$$

(We repeat here the earlier statement of this result.)

5.8.10 Theorem. *The following* ***reduced nonholonomic Lagrange-d'Alembert equations*** *or* ***Lagrange-d'Alembert-Poincaré equations*** *hold for each $1 \le \alpha \le \sigma$ and $1 \le b \le m$:*

$$\begin{aligned}
\frac{d}{dt}\frac{\partial l_c}{\partial \dot{r}^\alpha} - \frac{\partial l_c}{\partial r^\alpha} &= -\frac{\partial I^{cd}}{\partial r^\alpha} p_c p_d - \mathcal{D}^c_{b\alpha} I^{bd} p_c p_d - \mathcal{B}^c_{\alpha\beta} p_c \dot{r}^\beta \\
&\quad - \mathcal{D}_{\beta\alpha b} I^{bc} p_c \dot{r}^\beta - \mathcal{K}_{\alpha\beta\gamma} \dot{r}^\beta \dot{r}^\gamma, \\
\frac{d}{dt} p_b &= C^c_{ab} I^{ad} p_c p_d + \mathcal{D}^c_{b\alpha} p_c \dot{r}^\alpha + \mathcal{D}_{\alpha\beta b} \dot{r}^\alpha \dot{r}^\beta.
\end{aligned}$$

Here $l_c(r^\alpha, \dot{r}^\alpha, p_a)$ is the constrained Lagrangian; r^α, $1 \le \alpha \le \sigma$, are coordinates in the shape space; p_a, $1 \le a \le m$, are components of the momentum *map in the body representation,* $p_a = \langle \partial l_c / \partial \Omega_{\rm loc}, e_a(r) \rangle$*;* I^{ad} *are the components of the inverse locked inertia tensor;* $\mathcal{B}^a_{\alpha\beta}$ *are the local coordinates of the curvature* $\mathcal{B}$ *of the nonholonomic connection* $\mathcal{A}$*; and the coefficients* $\mathcal{D}^c_{b\alpha}$*,* $\mathcal{D}_{\alpha\beta b}$*,* $\mathcal{K}_{\alpha\beta\gamma}$ *are given by the formulae*

$$\begin{aligned}
\mathcal{D}^c_{b\alpha} &= \sum_{a'=1}^{k} -C^c_{a'b}\mathcal{A}^{a'}_\alpha + \gamma^c_{b\alpha} + \sum_{a'=m+1}^{k} \lambda_{a'\alpha} C^{a'}_{ab} I^{ac},\\
\mathcal{D}_{\alpha\beta b} &= \sum_{a'=m+1}^{k} \lambda_{a'\alpha}\left(\gamma^{a'}_{b\beta} - \sum_{b'=1}^{k} C^{a'}_{b'b}\mathcal{A}^{b'}_\beta\right),\\
\mathcal{K}_{\alpha\beta\gamma} &= \sum_{a'=1}^{k} \lambda_{a'\gamma}\mathcal{B}^{a'}_{\alpha\beta},
\end{aligned}$$

where

$$\lambda_{a'\alpha} = l_{a'\alpha} - \sum_{b'=1}^{k} l_{a'b'}\mathcal{A}^{b'}_\alpha := \frac{\partial l}{\partial \xi^{a'} \partial \dot{r}^\alpha} - \sum_{b'=1}^{k} \frac{\partial l}{\partial \xi^{a'} \partial \xi^{b'}}\mathcal{A}^{b'}_\alpha$$

for $a' = m+1, \ldots, k$*. Here* $C^{b'}_{a'c'}$ *are the structure constants of the Lie algebra defined by* $[e_{a'}, e_{c'}] = C^{b'}_{a'c'} e_{b'}$*,* $a', b', c' = 1, \ldots, k$*; and the coefficients* $\gamma^{c'}_{b\alpha}$ *are defined by*

$$\frac{\partial e_b}{\partial r^\alpha} = \sum_{c'=1}^{k} \gamma^{c'}_{b\alpha} e_{c'}.$$

Proof. The second set of equations, which are the momentum equations, were derived in the preceding proposition. To get the first set of equations, one can proceed in three ways. First, one can invoke the calculations for the motion relative to a general Ehresmann connection, restricting oneself to the variations that are horizontal; this is a straightforward, although somewhat tedious calculation. Alternatively, one can make use of the horizontal part of the calculations in Marsden and Scheurle [1993b], which as we remarked, are valid for any choice of connection. In particular, one can use the nonholonomic connection. A third method is to write the equations in a "vector" form similar to those for the momentum equation in body representation that we derived earlier by using the local form of the equations regarding the momentum terms as affine constraints (see equation (5.2.9) of the text):

$$\begin{aligned}
-\delta l_c = &- \left\langle \frac{\partial l}{\partial \xi}, d\mathcal{A}_{\rm loc}(\dot{r}, \delta r) - [\mathcal{A}_{\rm loc}(\dot{r}), \mathcal{A}_{\rm loc}(\delta r)] \right\rangle \\
&- \left\langle \frac{\partial l}{\partial \xi}, (D\mathbb{I}^{-1}_{\rm loc} p)(\delta r) \right\rangle. \qquad (5.8.11)
\end{aligned}$$

When these equations are converted to coordinate form and the basic dynamical variables are taken to be $(r, \dot{r}, \Omega = \mathbb{I}_{\mathrm{loc}}^{-1} p)$, one recovers the coordinate form above. ■

The above equations become the reduced Euler-Lagrange equations, also called the Lagrange-Poincaré equations in case there are no constraints. Notice also that the reduced equations are decoupled from the group variables, which is important for the reconstruction process. We summarize what we have already established on reconstruction as follows:

5.8.11 Proposition. *The group variables are reconstructed by means of the equation*

$$\dot{g} = g \cdot \xi$$

where $\xi = \Omega - \mathcal{A}_{\mathrm{loc}} \dot{r}$.

Of course we could also write this equation in terms of the nonholonomic momentum p_b. As before, let $\mathcal{A} : TQ \to \mathfrak{g}$ be the Lie algebra valued one-form corresponding to $A_q : T_q Q \to T_q(\mathrm{Orb}(q))$. Since the nonholonomic momentum map is equivariant, we can write it in a local trivialization, as before:

$$J^{\mathrm{nhc}}(g, r, \dot{g}, \dot{r}) = \mathrm{Ad}^*_{g^{-1}}(J^{\mathrm{nhc}}_{\mathrm{loc}}(r, \dot{r}, \xi)).$$

This is a form similar to that for the local expression for a connection and its curvature. Then the reconstruction equation becomes

$$\dot{g} = g(-\mathcal{A}_{\mathrm{loc}}(r)\dot{r} + \mathbb{I}_{\mathrm{loc}}^{-1}(r)p)$$

where $\mathcal{A}_{\mathrm{loc}} : T(Q/G) \to \mathfrak{g}$ is the local version of $\mathcal{A}$ and $\mathbb{I}_{\mathrm{loc}}^{-1}$ is the local version of the locked inertia tensor, as was defined before.

Note that $\dot{g}$ depends linearly on $\dot{r}$ and also linearly on p. In the case of horizontal symmetries, the term $-\mathcal{A}_{\mathrm{loc}}\dot{r}$ defines the *geometric phase* and the term $\Omega_{\mathrm{loc}} = \mathbb{I}_{\mathrm{loc}}^{-1}(r)p := \Gamma(r)p$ determines the *dynamic phase*. We adopt the same terminology in the general case. If the dynamic phase term is zero then the motion in the group variables is determined solely by the *path* in the base space, not its time parametrization. On the other hand, the dynamic phase determines the motion of the system when $\dot{r} = 0$ and hence corresponds to unforced motions of the system. For a system with horizontal symmetries, p is a constant.

As we have shown, it is possible to choose a basis of sections for $S_q = \mathcal{D}_q \cap T_q\mathrm{Orb}$ such that the momentum map and the locked inertia tensor is group invariant (independent of g). This was also shown by Ostrowski, Burdick, Lewis, and Murray [1995] , who write the momentum and reconstruction equations in the form

$$\begin{aligned} \dot{g} &= g(-\mathcal{A}_{\mathrm{loc}}(r)\dot{r} + \mathbb{I}(r)^{-1}p) \\ \dot{p} &= \sigma(r, \dot{r}, p) \end{aligned}$$

To reiterate, the reconstruction process now decouples as follows: given an initial condition and a path in the base space, we first integrate the momentum equation to determine $p(t)$ for all time. We then use $r(t)$ and $p(t)$ to determine the motion in the fiber. This decoupling is only possible when $\dot{p}$ is independent of g, since otherwise the equations for p and g are coupled. Of course, this whole process can be read in many different ways depending on the dynamics and control objectives.

The intrinsic geometry of the Lagrange-d'Alembert-Poincare equations is studied in detail in Cendra, Marsden, and Ratiu [2001b]

5.9 The Lagrangian and Hamiltonian Comparison.[1]

This section compares the Hamiltonian approach to systems with nonholonomic constraints (see Weber [1986], Arnold, Kozlov, and Neishtadt [1988], and Bates and Sniatycki [1993], van der Schaft and Maschke [1994] and references therein) with the Lagrangian approach (see Koiller [1992], Ostrowski [1995] and Bloch, Krishnaprasad, Marsden, and Ratiu [1996]). There are many differences in the approaches and each has its own advantages; some structures have been discovered on one side and their analogues on the other side are interesting to clarify. For example, the momentum equation and the reconstruction equation were first found on the Lagrangian side and are useful for the control theory of these systems, while the failure of the reduced two form to be closed (i.e., the failure of the Poisson bracket to satisfy the Jacobi identity) was first noticed on the Hamiltonian side. Clarifying the relation between these approaches is important for the future development of the control theory and stability and bifurcation theory for such systems. In addition to this work, we treat, in this unified framework, a simplified model of the bicycle (see Getz and Marsden [1995]), which is an important underactuated (nonminimum phase) control system.

Review of the Hamiltonian Formulation. Bates and Sniatycki [1993], hereafter denoted [BS], developed the Hamiltonian side, while Bloch, Krishnaprasad, Marsden, and Murray [1996], hereafter denoted [BKMM], explored the Lagrangian side and that approach has been reviewed in the preceding sections. It was not obvious how these two approaches were equivalent because, for example, [BKMM] developed the momentum equation and the reduced Lagrange-d'Alembert equations and it is not obvious how these correspond to the developments in [BS]. Our aim is to establish links between these two sides and use the ideas and results of each to shed light

[1] This section is based on Koon and Marsden [1997b]

on the other, with the goal of deepening our understanding of both points of view.

We illustrate the basic theory with the snakeboard, the well known example treated in [BKMM].

The approach of [BS] starts on the Lagrangian side with a configuration space Q and a Lagrangian L of the form kinetic energy minus potential energy, *i.e.,*

$$L(q, \dot{q}) = \frac{1}{2}\langle\langle \dot{q}, \dot{q}\rangle\rangle - V(q),$$

where $\langle\langle\ ,\ \rangle\rangle$ is a metric on Q defining the kinetic energy and V is a potential energy function. We do not restrict ourselves to Lagrangians of this form.

As above, our nonholonomic constraints are given by a distribution $\mathcal{D} \subset TQ$. We also let $\mathcal{D}^o \subset T^*Q$ denote the annihilator of this distribution.

As above, the basic equations are given by the Lagrange-d'Alembert principle.

The Legendre transformation $\mathbb{F}L : TQ \to T^*Q$, assuming that it is a diffeomorphism, is used to define the Hamiltonian $H : T^*Q \to \mathbb{R}$ in the standard fashion (ignoring the constraints for the moment):

$$H = \langle p, \dot{q}\rangle - L = p_i \dot{q}^i - L.$$

Here, the momentum is $p = \mathbb{F}L(v_q) = \partial L/\partial \dot{q}$. Under this change of variables, the equations of motion are written in the Hamiltonian form as

$$\dot{q}^i = \frac{\partial H}{\partial p_i}.$$
$$\dot{p}_i = -\frac{\partial H}{\partial q^i} + \lambda_a \omega_i^a,$$

where $i = 1, \ldots, n$, together with the constraint equations.

The preceding Hamiltonian equations can be rewritten as

$$i_X \Omega = dH + \lambda_a \pi_Q^* \omega^a, \tag{5.9.1}$$

where X is the vector field on T^*Q governing the dynamics, Ω is the canonical symplectic form on T^*Q, and $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection. We may write X in coordinates as $X = \dot{q}^i \partial_{q^i} + \dot{p}_i \partial_{p_i}$.

On Lagrangian side, we saw that one can get rid of the Lagrangian multipliers. On the Hamiltonian side, it is also desirable to model the Hamiltonian equations without the Lagrange multipliers by a vector field on a submanifold of T^*Q. We do this in what follows.

First of all, we define the set $\mathcal{M} = \mathbb{F}L(\mathcal{D}) \subset T^*Q$, so that the constraints on the Hamiltonian side are given by $p \in \mathcal{M}$. Besides $\mathcal{M}$, another basic object we deal with is defined as

$$\mathcal{F} = (T\pi_Q)^{-1}(\mathcal{D}) \subset TT^*Q.$$

Using a basis ω^a of the annihilator $\mathcal{D}^{\mathrm{o}}$, we can write these spaces as

$$\mathcal{M} = \{p \in T^*Q \mid \omega^a((\mathbb{F}L)^{-1}(p)) = 0\}, \tag{5.9.2}$$

and

$$\mathcal{F} = \{u \in TT^*Q \mid \left\langle \pi_Q^* \omega^a, u \right\rangle = 0\}. \tag{5.9.3}$$

Finally, we define

$$\mathcal{H} = \mathcal{F} \cap T\mathcal{M}.$$

Using natural coordinates $(q^i, p_i, \dot{q}^i, \dot{p}_i)$ on TT^*Q, we see that the distribution $\mathcal{F}$ naturally lifts the constraint on $\dot{q}$ from TQ to TT^*Q. On the other hand, the space $\mathcal{M}$ puts the associated constraints on the variable p and therefore the intersection $\mathcal{H}$ puts the constraints on both variables.

To eliminate the Lagrange multipliers, we regard the Hamiltonian equations as a vector field on the constraint submanifold $\mathcal{M} \subset T^*Q$ which takes values in the constraint distribution $\mathcal{H}$. Next we recall from [BS] how to construct these equations intrinsically using the ideas of symplectic geometry.

A result of [BS] is that $\Omega_{\mathcal{H}}$, the restriction of the canonical two-form Ω of T^*Q fiberwise to the distribution $\mathcal{H}$ of the constraint submanifold $\mathcal{M}$, is nondegenerate. Note that $\Omega_{\mathcal{H}}$ is not a true two form on a manifold, so it does not make sense to speak about it being closed. We speak of it as a fiber-restricted two form to avoid any confusion. Of course it still makes sense to talk about it being nondegenerate; it just means nondegenerate as a bilinear form on each fiber of $\mathcal{H}$. The dynamics is then given by the vector field $X_{\mathcal{H}}$ on $\mathcal{M}$ which takes values in the constraint distribution $\mathcal{H}$ and is determined by the condition

$$i_{X_{\mathcal{H}}} \Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{5.9.4}$$

where $dH_{\mathcal{H}}$ is the restriction of $dH_{\mathcal{M}}$ to $\mathcal{H}$. We will be exploring the coordinate meaning of this condition and its comparison with the Lagrangian formulation in the subsequent sections.

Lagrangian Side. We now construct the geometric structures on the tangent bundle TQ corresponding to those on the Hamiltonian side from the preceding subsection and formulate a similar procedure for obtaining the equations of motion. By doing this, it will be easier to made comparison with the geometric constructions and analytic formulations in [BKMM].

First of all, we can define the energy function E simply as $E = H \circ \mathbb{F}L$ and pull back to TQ the canonical two-form on T^*Q and denote it by Ω_L.

We define the distribution $\mathcal{C} = (T\tau_Q)^{-1}(\mathcal{D}) \subset TTQ$, where $\tau_Q : TQ \to Q$. In coordinates, the distribution $\mathcal{C}$ consists of vectors annihilated by the form $\tau_Q^* \omega^a$:

$$\mathcal{C} = \{u \in TTQ \mid \left\langle \tau_Q^* \omega^a, u \right\rangle = 0\}. \tag{5.9.5}$$

When $\mathcal{C}$ is restricted to the constraint submanifold $\mathcal{D} \subset TQ$, we obtain the constraint distribution $\mathcal{K}$:

$$\mathcal{K} = \mathcal{C} \cap T\mathcal{D}. \tag{5.9.6}$$

Clearly $\mathcal{M} = \mathbb{F}L(\mathcal{D})$ and $\mathcal{H} = T\mathbb{F}L(\mathcal{K})$.

The dynamics is given by a vector field $X_\mathcal{K}$ on the manifold $\mathcal{D}$ which takes values in $\mathcal{K}$ and satisfies the equation

$$i_{X_\mathcal{K}}\Omega_\mathcal{K} = dE_\mathcal{K}, \tag{5.9.7}$$

where $dE_\mathcal{K}$ and $\Omega_\mathcal{K}$ are the restrictions of $dE\mathcal{D}$ and $\Omega_\mathcal{D}$ respectively to the distribution $\mathcal{K}$ and where $E_\mathcal{D}$ and $\Omega_\mathcal{D}$ are the restrictions of E and Ω_L to $\mathcal{D}$.

The Equivalence of the Hamiltonian and the Lagrange-d'Alembert Formulations. The Lagrangian procedure on TQ formulated in the preceding subsection acts as a bridge between [BS] and [BKMM]. We can show the correctness of the Lagrangian procedure given above by (carefully) invoking the results of [BS] (generalized to arbitrary Lagrangians and with some gaps filled in), or by checking the methods against the results of [BKMM]. We choose the latter method.

5.9.1 Theorem. *Consider a configuration space Q, a hyperregular Lagrangian L and a distribution $\mathcal{D}$ that describes the kinematic nonholonomic constraints. The $\mathcal{K}$-valued vector field $X_\mathcal{K}$ on $\mathcal{D}$ given by the equation*

$$i_{X_\mathcal{K}}\Omega_\mathcal{K} = dE_\mathcal{K} \tag{5.9.8}$$

defines dynamics that are equivalent to the Lagrange-d'Alembert equations together with the constraints.

Proof. Consider the following form of the equations: $i_{X_\mathcal{H}}\Omega_\mathcal{M} = dH_\mathcal{M}$ on $\mathcal{H}$; that is,

$$\langle i_{X_\mathcal{H}}\Omega_\mathcal{M}, u\rangle = \langle dH_\mathcal{M}, u\rangle\,,$$

for all $u \in \mathcal{H}$. If we rewrite this in the form $\langle dH_\mathcal{M} - i_{X_\mathcal{H}}\Omega_\mathcal{M}, u\rangle = 0$, then on the Lagrangian side, this is nothing but

$$\langle dE_\mathcal{D} - i_{X_\mathcal{K}}(\Omega_L)_\mathcal{D}, v\rangle = 0,$$

where $v \in \mathcal{K}$. With appropriate interpretations, this is equivalent to Lagrange-d'Alembert principle:

$$\left(\frac{d}{dt}\frac{\partial L}{\partial \dot q^i} - \frac{\partial L}{\partial q^i}\right)(\delta q^i) = 0$$
$$\omega^a(\dot q) = 0$$

where $\omega(\delta q) = 0$. ■

Example: The Snakeboard. As in §5.8, the snakeboard is a modified version of a skateboard in which the front and back pairs of wheels are independently actuated. The extra degree of freedom enables the rider to generate forward motion by twisting their body back and forth, while simultaneously moving the wheels with the proper phase relationship. For details, see [BKMM] and the references listed there. Here we shall include some of the computations shown in that paper both for completeness as well as to make concrete the nonholonomic theory.

The snakeboard is modeled as a rigid body (the board) with two sets of independently actuated wheels, one on each end of the board. The human rider is modeled as a momentum wheel which sits in the middle of the board and is allowed to spin about the vertical axis. Spinning the momentum wheel causes a counter-torque to be exerted on the board. The configuration of the board is given by the position and orientation of the board in the plane, the angle of the momentum wheel, and the angles of the back and front wheels. Let (x, y, θ) represent the position and orientation of the center of the board, ψ the angle of the momentum wheel relative to the board, and ϕ_1 and ϕ_2 the angles of the back and front wheels, also relative to the board. Take the distance between the center of the board and the wheels to be r. See figure 5.9.1.

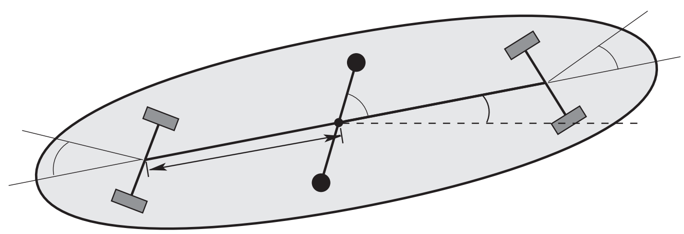

FIGURE 5.9.1. The geometry of the snakeboard.

In [BKMM], a simplification is made which we shall also assume here namely $\phi_1 = -\phi_2$, $J_1 = J_2$. The parameters are also chosen such that $J + J_0 + J_1 + J_2 = mr^2$, where m is the total mass of the board, J is the inertia of the board, J_0 is the inertia of the rotor and J_1, J_2 are the inertia of the wheels. This simplification eliminates some terms in the derivation but does not affect the essential geometry of the problem. Setting $\phi = \phi_1 = -\phi_2$, then the configuration space becomes $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ is the total kinetic energy of the system and is given by

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}J_0\dot{\psi}^2 + J_0\dot{\psi}\dot{\theta} + J_1\dot{\phi}^2.$$

The Constraints. The rolling of the front and rear wheels of the snakeboard is modeled using nonholonomic constraints which allow the wheels to spin about the vertical axis and roll in the direction that they are pointing. The wheels are not allowed to slide in the sideways direction. The constraints are defined by

$$-\sin(\theta+\phi)\dot{x}+\cos(\theta+\phi)\dot{y}-r\cos\phi\dot{\theta}=0 \tag{5.9.9}$$
$$-\sin(\theta-\phi)\dot{x}+\cos(\theta-\phi)\dot{y}+r\cos\phi\dot{\theta}=0 \tag{5.9.10}$$

and can be simplified to

$$\begin{aligned}\dot{x}&=-r\cot\phi\cos\theta\dot{\theta}\\ \dot{y}&=-r\cot\phi\sin\theta\dot{\theta}.\end{aligned}$$

Since the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ on the constraint submanifold $\mathcal{D}$ and its inverse are given by

$$\begin{aligned}p_x&=-mr\cot\phi\cos\theta\dot{\theta}\\ p_y&=-mr\cot\phi\sin\theta\dot{\theta}\\ p_\theta&=mr^2\dot{\theta}+J_0\dot{\psi}\\ p_\psi&=J_0\dot{\psi}+J_0\dot{\theta}\\ p_\phi&=2J_1\dot{\phi}\\ \dot{x}&=-\frac{r}{mr^2-J_0}\cot\phi\cos\theta(p_\theta-p_\psi)\\ \dot{y}&=-\frac{r}{mr^2-J_0}\cot\phi\sin\theta(p_\theta-p_\psi)\\ \dot{\theta}&=\frac{p_\theta-p_\psi}{mr^2-J_0}\\ \dot{\psi}&=\frac{mr^2p_\psi-J_0p_\theta}{J_0(mr^2-J_0)}\\ \dot{\phi}&=\frac{p_\phi}{2J_1},\end{aligned}$$

the constraint submanifold $\mathcal{M}$ is defined by

$$\begin{aligned}\mathcal{M}&=\{(x,y,\theta,\psi,\phi,p_x,p_y,p_\theta,p_\psi,p_\phi)\mid\\ p_x&=-\frac{mr}{mr^2-J_0}\cot\phi\cos\theta(p_\theta-p_\psi),p_y\\ &=-\frac{mr}{mr^2-J_0}\cot\phi\sin\theta(p_\theta-p_\psi).\}\end{aligned}$$

Notice that $\mathcal{M}$ may be thought of as a graph in T^*Q and we can use the induced coordinates $(x,y,\theta,\psi,\phi,p_\theta,p_\psi,p_\phi)$ as its local coordinates. Hence the distribution $\mathcal{H}$ of $\mathcal{M}$ is

$$\begin{aligned}\mathcal{H}&=\ker\{dx+r\cot\phi\cos\theta d\theta,dy+r\cot\phi\sin\theta d\theta\}\\ &=\operatorname{span}\{-r\cot\phi\cos\theta\partial_x-r\cot\phi\sin\theta\partial_y+\partial_\theta,\partial_\psi,\partial_\phi,\partial_{p_\theta},\partial_{p_\psi},\partial_{p_\phi}\}.\end{aligned}$$

The Hamiltonian. The corresponding Hamiltonian is given via the Legendre transform by

$$H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{1}{2J_0}p_\psi^2 + \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1}p_\phi^2.$$

Now if we restrict the Hamiltonian H to the submanifold $\mathcal{M}$, we get

$$\begin{aligned} H_{\mathcal{M}} =& \frac{mr^2}{2(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi)^2 + \frac{1}{2J_0}p_\psi^2 \\ &+ \frac{1}{2(mr^2 - J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1}p_\phi^2. \end{aligned}$$

After computing its differential $dH_{\mathcal{M}}$ and restricting it to $\mathcal{H}$, we have

$$\begin{aligned} dH_{\mathcal{H}} = &-\frac{mr^2}{(mr^2 - J_0)^2}\cot\phi\csc^2\phi(p_\theta - p_\psi)^2 d\phi \\ &+ \frac{mr^2}{(mr^2 - J_0)^2}\cot^2\phi(p_\theta - p_\psi)(dp_\theta - dp_\psi) \\ &+ \frac{1}{J_0}p_\psi dp_\psi + \frac{1}{(mr^2 - J_0)}(p_\theta - p_\psi)(dp_\theta - dp_\psi) + \frac{1}{2J_1}p_\phi dp_\phi. \end{aligned}$$

The Two Form. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned} \Omega_{\mathcal{M}} &= dx \wedge dp_x + dy \wedge dp_y + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\ &= k dx \wedge [\csc^2\phi\cos\theta(p_\theta - p_\psi)d\phi + \cot\phi\sin\theta(p_\theta - p_\psi)d\theta \\ &\quad - \cot\phi\cos\theta(dp_\theta - dp_\psi)] \\ &\quad + k dy \wedge [\csc^2\phi\sin\theta(p_\theta - p_\psi)d\phi - \cot\phi\cos\theta(p_\theta - p_\psi)d\theta \\ &\quad - \cot\phi\sin\theta(dp_\theta - dp_\psi)] \\ &\quad + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi, \end{aligned}$$

where $k = mr/(mr^2 - J_0)$. If we restrict $\Omega_{\mathcal{M}}$ to the distribution $\mathcal{H}$, we get

$$\begin{aligned} \Omega_{\mathcal{H}} &= -kr\cot\phi\cos\theta d\theta \wedge [\csc^2\phi\cos\theta(p_\theta - p_\psi)d\phi \\ &\qquad + \cot\phi\sin\theta(p_\theta - p_\psi)d\theta - \cot\phi\cos\theta(dp_\theta - dp_\psi)] \\ &\quad - kr\cot\phi\sin\theta d\theta \wedge [\csc^2\phi\sin\theta(p_\theta - p_\psi)d\phi \\ &\qquad - \cot\phi\cos\theta(p_\theta - p_\psi)d\theta - \cot\phi\sin\theta(dp_\theta - dp_\psi)] \\ &\quad + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\ &= d\theta \wedge [-kr\cot\phi\csc^2\phi(p_\theta - p_\psi)d\phi \\ &\quad + kr\cot^2\phi(dp_\theta - dp_\psi) + dp_\theta] + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi. \end{aligned}$$

Equations of Motion. Notice that any vector field $X_{\mathcal{M}}$ is of the form

$$X_{\mathcal{M}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\theta\partial_{p_\theta} + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi}.$$

But $X_{\mathcal{H}}$ also lies in $\mathcal{H} = \ker\{dx + r\cot\phi\cos\theta d\theta, dy + r\cot\phi\sin\theta d\theta\}$ and hence must be of the form

$$\begin{aligned} X_{\mathcal{H}} = &\dot{\theta}(-r\cot\phi\cos\theta\partial_x - r\cot\phi\sin\theta\partial_y + \partial_\theta) \\ &+ \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\theta\partial_{p_\theta} + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi}, \end{aligned}$$

which gives us the first set of relationships

$$\begin{aligned} \dot{x} &= -r\cot\phi\cos\theta\,\dot{\theta} \\ \dot{y} &= -r\cot\phi\sin\theta\,\dot{\theta}. \end{aligned}$$

Moreover,

$$\begin{aligned} i_{X_{\mathcal{H}}}\Omega_{\mathcal{H}} = &- kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot{\theta}\,d\phi + kr\cot^2\phi\dot{\theta}(dp_\theta - dp_\psi) + \dot{\theta}\,dp_\theta \\ &+ \dot{\psi}\,dp_\psi + kr\cot\phi\csc^2\phi(p_\theta - p_\psi)\dot{\phi}\,d\theta + \dot{\phi}\,dp_\phi - kr\cot^2\phi\dot{p}_\theta\,d\theta \\ &-\dot{p}_\theta\,d\theta + \; kr\cot^2\phi\dot{p}_\psi\,d\theta - \dot{p}_\psi\,d\psi - \dot{p}_\phi\,d\phi, \end{aligned}$$

and if equated with $dH_{\mathcal{H}}$ and after simplification, we have

$$\dot{p}_\theta = \frac{\cot\phi}{2J_1(1 - \frac{J_0}{mr^2}\sin^2\phi)}p_\phi(p_\theta - p_\psi) \tag{5.9.11}$$

$$\dot{p}_\psi = 0 \tag{5.9.12}$$

$$\dot{p}_\phi = 0 \tag{5.9.13}$$

$$\dot{\theta} = \frac{p_\theta - p_\psi}{mr^2 - J_0} \tag{5.9.14}$$

$$\dot{\psi} = \frac{mr^2 p_\psi - J_0 p_\theta}{J_0(mr^2 - J_0)} \tag{5.9.15}$$

$$\dot{\phi} = \frac{p_\phi}{2J_1}. \tag{5.9.16}$$

Notice that the last 3 equations are nothing but the inverse of the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$ written in local coordinates. The first equation is equivalent to the momentum equation (discussed below and in [BKMM]) written in Hamiltonian form and the 2nd and 3rd equations are the reduced equations on the shape space, again in their Hamiltonian forms.

Moreover, the corresponding Lagrangian procedure gives the equations of the motion on the Lagrangian side as

$$\ddot{\theta} - \cot\phi\dot{\phi}\dot{\theta} + \frac{J_0}{mr^2}\sin^2\phi\ddot{\psi} = 0 \tag{5.9.17}$$

$$J_0\ddot{\psi} + J_0\ddot{\theta} = 0 \tag{5.9.18}$$

$$2J_1\ddot{\phi} = 0 \tag{5.9.19}$$

and it can be shown that both systems of equations are equivalent via the Legendre transform $\mathbb{F}L|\mathcal{D}$.

Nonholonomic Mechanical Systems with Symmetry Now we add the hypothesis of symmetry to the preceding development. We have already studied this situation from the Lagrangian side. Recall that the momentum equation extends the Noether Theorem to nonholonomic systems by deriving an equation for the momentum map that replace the usual conservation law. Recall that if the Lagrangian L is invariant under the group action and that ξ^q is a section of the bundle $\mathfrak{g}^{\mathcal{D}}$, then any solution $q(t)$ of the Lagrange-d'Alembert equations must satisfy, in addition to the given kinematic constraints, the momentum equation:

$$\frac{d}{dt}\left(J^{\mathrm{nh}}(\xi^{q(t)})\right) = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}(\xi^{q(t)})\right]_Q^i. \tag{5.9.20}$$

When the momentum map is paired with a section in this way, we will just refer to it as the momentum. As we have seen earlier the nonholonomic momentum map may or may not be conserved. Recall also that this equation can be conveniently written in body representation.

Recall that with the help of the nonholonomic mechanical connection, the Lagrange-d'Alembert principle may be broken up into two principles by breaking the variations δq into two parts, namely parts that are horizontal with respect to the nonholonomic connection and parts that are vertical (but still in $\mathcal{D}$), and the reduced equations break up into *two* sets: a set of reduced Lagrange-d'Alembert equations (which have curvature terms appearing as 'forcing'), and a momentum equation, which have a form generalizing the components of the Euler-Poincaré equations along the symmetry directions consistent with the constraints. When one supplements these equations with the reconstruction equations, one recovers the full set of equations of motion for the system.

Hamiltonian Reduction. In working out the nonholonomic Hamiltonian reduction, [BS] also starts out with a *simple nonholonomic mechanical system*. Recall that the Legendre transformation $\mathbb{F}L : TQ \to T^*Q$ is used to define the constraint submanifold $\mathcal{M} \subset T^*Q$ where

$$\mathcal{M} = \mathbb{F}L(\mathcal{D}). \tag{5.9.21}$$

On this manifold, there is a distribution $\mathcal{H}$

$$\mathcal{H} = \mathcal{F} \cap T\mathcal{M}, \tag{5.9.22}$$

where

$$\mathcal{F} = (T\pi)^{-1}(\mathcal{D}), \tag{5.9.23}$$

and $\pi : T^*Q \to Q$. Also recall that $\Omega_{\mathcal{H}}$, the restriction of the canonical two-form Ω of T^*Q to the distribution $\mathcal{H}$ of the constraint submanifold $\mathcal{M}$, is nondegenerate and that the dynamics is given by a vector field $X_{\mathcal{H}}$ on $\mathcal{M}$ taking values in $\mathcal{H}$ and satisfies the equation

$$i_{X_{\mathcal{H}}}\Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{5.9.24}$$

where $dH_{\mathcal{H}}$ is the (fiberwise) restriction of $dH_{\mathcal{M}}$ to $\mathcal{H}$.

Now let G be the symmetry group of this system and assume that the quotient space $\overline{\mathcal{M}} = \mathcal{M}/G$ of the G-orbit in $\mathcal{M}$ is a quotient manifold with projection map $\rho : \mathcal{M} \to \overline{\mathcal{M}}$. Since G is a symmetry group, all intrinsically defined vector fields and distributions push down to $\overline{\mathcal{M}}$. In particular, the vector field $X_{\mathcal{M}}$ on $\mathcal{M}$ pushes down to a vector field $\overline{X}_{\overline{\mathcal{M}}} = \rho_* X_{\mathcal{M}}$, and the distribution $\mathcal{H}$ pushes down to a distribution $\rho_*\mathcal{H}$ on $\overline{\mathcal{M}}$.

However, $\Omega_{\mathcal{H}}$ need not push down to a two-form defined on $\rho_*\mathcal{H}$, despite the fact that $\Omega_{\mathcal{H}}$ is G-invariant. This is because there may be infinitesimal symmetry $\xi_{\mathcal{M}}$ that lies in $\mathcal{H}$ such that $i_{\xi_{\mathcal{M}}}\Omega_{\mathcal{H}} \neq 0$, To eliminate this difficulty, [BS] restricts $\Omega_{\mathcal{H}}$ to a subdistribution $\mathcal{U}$ of $\mathcal{H}$ defined by

$$\mathcal{U} = \{u \in \mathcal{H} \mid \Omega_{\mathcal{H}}(u, v) = 0 \quad \text{for all} \quad v \in \mathcal{V}\cap\mathcal{H}\} = \mathcal{H}\cap(\mathcal{V}\cap\mathcal{H})^{\perp}, \tag{5.9.25}$$

where $\mathcal{V}$ is the distribution on $\mathcal{M}$ tangent to the orbits of G in $\mathcal{M}$ and is spanned by the infinitesimal symmetries and $(\mathcal{V}\cap\mathcal{H})^{\perp}$ is the $\Omega_{\mathcal{H}}$-orthogonal complement of $(\mathcal{V}\cap\mathcal{H})$. Clearly, $\mathcal{U}$ and $\mathcal{V}$ are both G-invariant, project down to $\overline{\mathcal{M}}$ and $\rho_*\mathcal{V} = 0$. Define $\overline{\mathcal{H}}$ by

$$\overline{\mathcal{H}} = \rho_*\mathcal{U}. \tag{5.9.26}$$

It is proven in [BS] that

1. The vector field $X_{\mathcal{H}}$ which satisfies the above Hamiltonian equation of motion (5.9.24) lies in the distribution $\mathcal{U}$.

2. The restriction $\Omega_{\mathcal{U}}$ of Ω to the distribution $\mathcal{U}$ pushes down to a nondegenerate 2-form $\Omega_{\overline{\mathcal{H}}} = \rho_*\Omega_{\mathcal{U}}$ on $\overline{\mathcal{H}}$, which is modeled by the symplectic space $(\mathcal{V} \cap \mathcal{H})^{\perp}/(\mathcal{V} \cap \mathcal{H}) \cap (\mathcal{V} \cap \mathcal{H})^{\perp}$.

3. Furthermore,
$$i_{\overline{X}_{\overline{\mathcal{H}}}}\Omega_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{H}}}, \tag{5.9.27}$$
where $h_{\overline{\mathcal{M}}} = \rho_* H_{\mathcal{M}}$ is the pushdown of the restriction to $\mathcal{M}$ of the Hamiltonian H and $dh_{\overline{\mathcal{H}}}$ is the restriction of $dh_{\overline{\mathcal{M}}}$ to $\overline{\mathcal{H}}$. This is because the equation $i_{X_{\mathcal{H}}}\Omega_{\mathcal{H}} = dH_{\mathcal{H}}$, restricted to $\mathcal{U} \subset \mathcal{H}$, vanishes on vectors in $\mathcal{V}$, and is G-invariant. Hence both sides push down to $\overline{\mathcal{H}}$.

Note that the original equations of motion are

$$i_{X_{\mathcal{H}}}\Omega_{\mathcal{H}} = dH_{\mathcal{H}} \tag{5.9.28}$$

where $\mathcal{H}$ is a distribution in the constraint manifold $\mathcal{M}$. After the reduction of symmetry we obtain equations of the same type

$$i_{\overline{X}_{\overline{\mathcal{H}}}}\Omega_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{H}}}, \tag{5.9.29}$$

where $\overline{\mathcal{H}}$ is a distribution in the reduced space $\overline{\mathcal{M}} = \mathcal{M}/G$.

Lagrangian Side. By using the Legendre transformation $\mathbb{F}L$, we can construct dual geometric structures on the tangent bundle TQ and formulate a similar Lagrangian reduction procedure. This allows us to better compare with the geometric constructions and analytic formulations on the manifold Q in [BKMM], and in the course of doing this, we realize that the requirement that the vector field $X_{\mathcal{H}}$ lies in the subdistribution $\mathcal{U}$ is equivalent to the extended Noether Theorem; that is, that any solution of the Lagrange-d'Alembert equations must satisfy the momentum equation.

We consider $\mathcal{D}$ as a constraint submanifold of TQ and then construct the distribution

$$\mathcal{K} = \mathcal{C} \cap T\mathcal{D}, \tag{5.9.30}$$

on TTQ, where

$$\mathcal{C} = (T\tau_Q)^{-1}(\mathcal{D}), \tag{5.9.31}$$

and $\tau_Q : TQ \to Q$. Clearly $\mathcal{D} = (\mathbb{F}L)^{-1}(\mathcal{M}), \mathcal{K} = (T\mathbb{F}L)^{-1}(\mathcal{H})$. The motion is then given by a vector field $X_{\mathcal{K}}$ on the manifold $\mathcal{D}$ which takes values in $\mathcal{K}$ and satisfies the equation

$$i_{X_{\mathcal{K}}}\Omega_{\mathcal{K}} = dE_{\mathcal{K}}, \tag{5.9.32}$$

where $dE_{\mathcal{K}}$ and $\Omega_{\mathcal{K}}$ are the restrictions of $dE_{\mathcal{D}}$ and $\Omega_{\mathcal{D}}$ respectively to the distribution $\mathcal{K}$.

Now let G be the symmetry group of this system and assume that the quotient space $\overline{\mathcal{D}} = \mathcal{D}/G$ of the G-orbit in $\mathcal{D}$ is a smooth quotient manifold with projection map $\lambda : \mathcal{D} \to \overline{\mathcal{D}}$. Since G is a symmetry group, all intrinsically defined vector fields and distributions push down to $\overline{\mathcal{D}}$. In particular, the vector field $X_{\mathcal{D}}$ on $\mathcal{D}$ pushes down to a vector field $\overline{X}_{\overline{\mathcal{D}}} = \lambda_* X_{\mathcal{D}}$, and the distribution $\mathcal{K}$ pushes down to a distribution $\lambda_*\mathcal{K}$ on $\overline{\mathcal{D}}$. Here we use the push forward symbol λ_* to mean that the vector fields are λ-related.

For the same reason as the Hamiltonian side, $\Omega_{\mathcal{K}}$ need not push down to a two-form defined on $\lambda_*\mathcal{K}$, despite the fact that $\Omega_{\mathcal{K}}$ is G-invariant. We can restrict $\Omega_{\mathcal{K}}$ to the subdistribution $\mathcal{W}$ of $\mathcal{K}$ defined by

$$\mathcal{W} = \{w \in \mathcal{K} \mid \Omega_{\mathcal{K}}(w, v) = 0 \text{ for all } v \in \mathcal{T} \cap \mathcal{K}\} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{D})^{\perp}, \tag{5.9.33}$$

where $\mathcal{T}$ is the distribution on $\mathcal{D}$ tangent to the orbits of G in $\mathcal{D}$ and is spanned by the infinitesimal symmetries. Clearly, $\mathcal{W}$ and $\mathcal{T}$ are both G-invariant, $\mathcal{W}$ projects down to $\overline{\mathcal{D}}$ and $\lambda_*\mathcal{T} = 0$. Define $\overline{\mathcal{K}}$ by

$$\overline{\mathcal{K}} = \lambda_*\mathcal{W}. \tag{5.9.34}$$

Since the above constructions are dual to those in the Hamiltonian side, we also have

1. The vector field $X_{\mathcal{K}}$ which satisfies the above equation (5.9.32) takes values in the distribution $\mathcal{W}$.

2. The restriction $\Omega_{\mathcal{W}}$ of Ω_L to the distribution $\mathcal{W}$, pushes down to a nondegenerate 2-form $\Omega_{\overline{\mathcal{K}}} = \lambda_* \Omega_{\mathcal{W}}$ on $\overline{\mathcal{K}}$, which is modeled by the symplectic space $(\mathcal{T} \cap \mathcal{K})^{\perp} / (\mathcal{T} \cap \mathcal{K}) \cap (\mathcal{T} \cap \mathcal{K})^{\perp}$.

3. The reduced equations of motion are given by
$$i_{\overline{X}_{\overline{\mathcal{K}}}} \Omega_{\overline{\mathcal{K}}} = d\overline{E}_{\overline{\mathcal{K}}}, \tag{5.9.35}$$
where $\overline{E}_{\overline{\mathcal{D}}} = \lambda_* E_{\mathcal{D}}$ is the pushdown of the restriction to $\mathcal{D}$ of the energy function E. This is because the equation $i_{X_{\mathcal{K}}} \Omega_{\mathcal{K}} = dE_{\mathcal{K}}$, restricted to $\mathcal{W} \subset \mathcal{K}$, vanishes on vectors in $\mathcal{T}$, and is G-invariant. Hence both sides push down to $\overline{\mathcal{K}}$. All these will become clearer in the subsequent computations.

The Equivalence of Hamiltonian and Lagrangian Reductions

5.9.2 Theorem. *Consider a simple nonholonomic mechanical system with symmetry and assume that it is in the principal case. Then the reduction procedure on TQ gives the constrained reduced nonholonomic equations of motion (as in* [BKMM]*).*

Proof The first difficulty is how to represent the constraint submanifold $\mathcal{D} \subset TQ$ in a way that is both intrinsic and ready for reduction. The comparison with the geometric constructions in [BKMM] and the desire to have the dynamics break up in a way that is ready for reconstruction give hints that we should use the tools like nonholonomic momentum p and the nonholonomic connection $\mathcal{A}$ in [BKMM] to describe the constraint submanifold $\mathcal{D}$

Recall that in [BKMM], the nonholonomic constraints together with the basic identity of the nonholonomic momentum map are used to synthesis a nonholonomic connection $\mathcal{A}$ and the nonholonomic constraints are then written in the form
$$g^{-1}\dot{g} = -A(r)\dot{r} + \Gamma(r)p, \tag{5.9.36}$$
where p is G-invariant. Hence, the constraint manifold is nothing but
$$\mathcal{D} = \{(g, r, \dot{g}, \dot{r}) \mid \dot{g} = g(-A(r)\dot{r} + \Gamma(r))p)\}. \tag{5.9.37}$$
It is a submanifold in TQ and we can use $(g, r, \dot{r}, p)$ as its induced local coordinates. Then, clearly, the corresponding coordinates for $\overline{\mathcal{D}} = \mathcal{D}/G$ are $(r, \dot{r}, p)$. From now on, we will use $A(r)$ to abbreviate $\mathcal{A}^{\mathrm{nh}}_{\mathrm{loc}}(r)$.

The next difficulty is to find the corresponding representations for the distribution $\mathcal{K}$, the subdistribution $\mathcal{T} \cap \mathcal{K}$ and its annihilator distribution $\mathcal{W}$ where
$$\mathcal{W} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{K})^{\perp}. \tag{5.9.38}$$
Recall that in [BKMM], a body fixed basis $e_b(g, r) = \mathrm{Ad}_g \cdot e_b(r)$ has been constructed such that the infinitesimal generators $(e_i(g, r))_Q$ of its first m

elements at a point q span $\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q))$. Assume that G is a matrix group and e_i^d is the component of $e_i(r)$ with respect to a fixed basis $\{b_a\}$ of the Lie algebra $\mathfrak{g}$ where $(b_a)_Q = \partial_{g^a}$, then

$$(e_i(g,r))_Q = g_d^a e_i^d \partial_{g^a}.$$

Since $\mathcal{K} = (T\tau)^{-1}(\mathcal{D})$ where $\mathcal{D}_q$ is the direct sum of $\mathcal{S}_q$ and the horizontal space of the nonholonomic connection $\mathcal{A}^{\mathrm{nh}}$, it can be represented in the induced coordinates by

$$\mathcal{K} = \mathrm{span}\{g_d^a e_i^d \partial_{g^a}, -g_b^a A_\alpha^b \partial_{g^a} + \partial_{r^\alpha}, \partial_{\dot{r}}, \partial_p\}. \tag{5.9.39}$$

Also, we have

$$\mathcal{T} \cap \mathcal{K} = \mathrm{span}\{g_d^a e_i^d \partial_{g^a}\}. \tag{5.9.40}$$

To find the distribution $\mathcal{W}$, we have to compute $i_{g_d^a e_i^d \partial_{g^a}} \Omega_{\mathcal{D}}$, for all $i = 1, \ldots, m$. Since L is G-invariant, we have

$$\begin{aligned}
\Omega_{\mathcal{D}} &= dg^a \wedge d\left(\frac{\partial L}{\partial \dot{g}^a}\right) + dr^\alpha \wedge d\left(\frac{\partial L}{\partial \dot{r}^\alpha}\right) \\
&= dg^a \wedge d\left((g^{-1})_a^b \frac{\partial l}{\partial \xi^b}\right) + dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right) \\
&= \frac{\partial (g^{-1})_a^b}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^a \wedge dg^c + (g^{-1})_a^b dg^a \wedge d\left(\frac{\partial l}{\partial \xi^b}\right) + dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot{r}^\alpha}\right).
\end{aligned}$$

Hence

$$\begin{aligned}
i_{(g_f^a e_i^f \partial_{g^a})} \Omega_{\mathcal{D}} &= g_f^a e_i^f \frac{\partial (g^{-1})_a^b}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^c - g_f^c e_i^f \frac{\partial (g^{-1})_a^b}{\partial g^c} \frac{\partial l}{\partial \xi^b} dg^a + e_i^b d\left(\frac{\partial l}{\partial \xi^b}\right) \\
&= e_i^f \left(\left(g_f^c \frac{\partial (g^{-1})_c^b}{\partial g^a} - \frac{\partial (g^{-1})_a^b}{\partial g^c} g_f^c\right) \frac{\partial l}{\partial \xi^b} dg^a + d\left(\frac{\partial l}{\partial \xi^f}\right)\right) \\
&= e_i^f \left((g^{-1})_\sigma^b \left(-\frac{\partial g_f^\sigma}{\partial g^\tau} g_a^\tau + \frac{\partial g_a^\sigma}{\partial g^\tau} g_f^\tau\right) \frac{\partial l}{\partial \xi^b} (g^{-1})_e^a dg^e + d\left(\frac{\partial l}{\partial \xi^f}\right)\right) \\
&= e_i^f \left(-C_{af}^b \frac{\partial l}{\partial \xi^b} (g^{-1})_e^a dg^e + d\left(\frac{\partial l}{\partial \xi^f}\right)\right) \\
&= dp_i - \frac{\partial l}{\partial \xi^f} d(e_i^f) - C_{af}^b \frac{\partial l}{\partial \xi^b} e_i^f (g^{-1})_e^a dg^e.
\end{aligned}$$

Here, C_{af}^b is the structural constants for the Lie algebra $\mathfrak{g}$ and

$$p_i = \frac{\partial l}{\partial \xi^f} e_i^f.$$

Therefore, the subdistribution $\mathcal{W} \subset \mathcal{K}$ is

$$\mathcal{W} = \ker\left\{dp_i - \frac{\partial l}{\partial \xi^f} d(e_i^f) - C_{af}^b \frac{\partial l}{\partial \xi^b} e_i^f (g^{-1})_e^a dg^e\right\}. \tag{5.9.41}$$

Since the constraint manifold $\mathcal{D}$ has the induced local coordinates $(g, r, \dot{r}, p)$, any vector field $X_{\mathcal{D}}$ on the manifold $\mathcal{D}$ is of the form

$$X_{\mathcal{D}} = \dot{g}^a \partial_{g^a} + \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i}.$$

If $X_{\mathcal{D}}$ lies in the distribution $\mathcal{K}$, then we have $\dot{g} = g(-A\dot{r} + \Gamma p)$. Moreover, if $X_{\mathcal{D}}$ lies in the distribution $\mathcal{W}$, then for each j, we have

$$\dot{p}_j - \frac{\partial l}{\partial \xi^d} \frac{\partial e_j^d}{\partial r^\alpha} \dot{r}^\alpha - C_{ad}^b \frac{\partial l}{\partial \xi^b} \xi^a e_j^d = 0, \qquad \text{i.e.,} \quad \dot{p}_j = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_j] + \dot{e}_j \right\rangle, \tag{5.9.42}$$

which gives the momentum equation. Therefore, any vector field $X_{\mathcal{W}}$ taking values in $\mathcal{W}$ must be of the form

$$X_{\mathcal{W}} = g_b^a \xi^b \partial_{g^a} + \dot{r}^\alpha \partial_{r^\alpha} + \ddot{r}^\alpha \partial_{\dot{r}^\alpha} + \dot{p}_i \partial_{p_i}, \tag{5.9.43}$$

where

$$\xi = -A\dot{r} + \Gamma p \qquad \dot{p}_j = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_j] + \dot{e}_j \right\rangle. \tag{5.9.44}$$

Now we are ready to do the reduction. But before that, we need to compute all the ingredients of the equation

$$i_{X_{\mathcal{K}}} \Omega_{\mathcal{K}} = dE_{\mathcal{K}}. \tag{5.9.45}$$

Notice first that since E is G-invariant, we have

$$\begin{aligned} E &= \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L = \frac{\partial L}{\partial \dot{g}^a} \dot{g}^a + \frac{\partial L}{\partial \dot{r}^\alpha} \dot{r}^\alpha - L \\ &= \frac{\partial l}{\partial \xi^a} \xi^a + \frac{\partial l}{\partial \dot{r}^\alpha} \dot{r}^\alpha - l. \end{aligned}$$

After restricting it to the submanifold $\mathcal{D}$, we have

$$\begin{aligned} E_{\mathcal{D}} &= \frac{\partial l}{\partial \xi^a} (-A_\alpha^a \dot{r}^\alpha + \Gamma^{ai} p_i) + \left(\frac{\partial l_c}{\partial \dot{r}^\alpha} + A_\alpha^a \frac{\partial l}{\partial \xi^a} \right) \dot{r}^\alpha - l_c \\ &= \frac{\partial l}{\partial \xi^a} \Gamma^{ai} p_i + \frac{\partial l_c}{\partial \dot{r}^\alpha} \dot{r}^\alpha - l_c. \end{aligned}$$

Therefore,

$$\begin{aligned} dE_{\mathcal{D}} = & \frac{\partial l}{\partial \xi^a} \left(\frac{\partial \Gamma^{ai}}{\partial r^\alpha} p_i dr^\alpha + \Gamma^{ai} dp_i \right) \\ & + \Gamma^{ai} p_i \left(\frac{\partial^2 l}{\partial r^\alpha \partial \xi^a} dr^\alpha + \frac{\partial^2 l}{\partial \dot{r}^\alpha \partial \xi^a} d\dot{r}^\alpha + \frac{\partial^2 l}{\partial p_j \partial \xi^a} dp_j \right) \\ + \ \dot{r}^\alpha & \left(\frac{\partial^2 l_c}{\partial r^\beta \partial \dot{r}^\alpha} dr^\beta + \frac{\partial^2 l_c}{\partial \dot{r}^\beta \partial \dot{r}^\alpha} d\dot{r}^\beta \frac{\partial^2 l_c}{\partial p_i \partial \dot{r}^\alpha} dp_i \right) - \frac{\partial l_c}{\partial r^\alpha} dr^\alpha - \frac{\partial l_c}{\partial p_i} dp_i. \end{aligned}$$

Furthermore,

$$
\begin{aligned}
&i_{X_{\mathcal{K}}}\Omega_{\mathcal{D}} \\
&\quad = g^a_f \xi^f \frac{\partial (g^{-1})^b_a}{\partial g^c}\frac{\partial l}{\partial \xi^b} dg^c - g^c_f \xi^f \frac{\partial (g^{-1})^b_a}{\partial g^c}\frac{\partial l}{\partial \xi^b} dg^a \\
&\qquad + g^a_f \xi^f (g^{-1})^b_a d\left(\frac{\partial l}{\partial \xi^b}\right) \\
&\qquad - \left(\frac{\partial}{\partial r^\alpha}\left(\frac{\partial l}{\partial \xi^b}\right)\dot r^\alpha + \frac{\partial}{\partial \dot r^\alpha}\left(\frac{\partial l}{\partial \xi^b}\right)\ddot r^\alpha + \frac{\partial}{\partial p_i}\left(\frac{\partial l}{\partial \xi^b}\right)\dot p_i\right)(g^{-1})^b_a dg^a \\
&\qquad + i_{(\dot r^\alpha \partial_{r^\alpha} + \ddot r^\alpha \partial_{\dot r^\alpha} + \dot p_i \partial p_i)}\left(dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot r^\alpha}\right)\right) \\
&\quad = \xi^f d\left(\frac{\partial l}{\partial \xi^f}\right) + \left(C^b_{fa}\frac{\partial l}{\partial \xi^b}\xi^f - \frac{d}{dt}\left(\frac{\partial l}{\partial \xi^a}\right)\right)(g^{-1})^a_e dg^e \\
&\qquad + \ i_{(\dot r^\alpha \partial_{r^\alpha} + \ddot r^\alpha \partial_{\dot r^\alpha} + \dot p_i \partial p_i)}\left(dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot r^\alpha}\right)\right).
\end{aligned} \tag{5.9.46}
$$

Clearly, both sides of the equation

$$i_{X_{\mathcal{K}}}\Omega_{\mathcal{K}} = \mathbf{d}E_{\mathcal{K}} \tag{5.9.47}$$

are G-invariant, and when restricted to the subdistribution $\mathcal{W} \subset \mathcal{K}$, they vanish on the distribution $\mathcal{T} \cap \mathcal{K}$. This can be shown to be true either by invoking how $\mathcal{W}$ has been constructed or by direct calculation, noticing that when

$$\left(C^b_{fa}\frac{\partial l}{\partial \xi^b}\xi^f - \frac{d}{dt}\left(\frac{\partial l}{\partial \xi^a}\right)\right)(g^{-1})^a_e dg^e \tag{5.9.48}$$

is paired with $g^f_c e^c_i$ in $\mathcal{T} \cap \mathcal{K}$, it is equal to zero on $\mathcal{W}$. Hence both sides push down to $\overline{\mathcal{K}}$ where

$$\overline{X}_{\overline{\mathcal{K}}} = \dot r^\alpha \partial_{r^\alpha} + \ddot r^\alpha \partial_{\dot r^\alpha} + \dot p_i \partial_{p_i}, \tag{5.9.49}$$

with

$$\dot p_i = \left\langle \frac{\partial l}{\partial \xi}, [\xi, e_i] + \dot e_i \right\rangle. \tag{5.9.50}$$

To find the remaining reduced equations, notice that the restriction of (5.9.48) to the subdistribution spanned by $\{-g^a_b A^b_\alpha \partial_{g^a} + \partial_{r^\alpha}, \partial_{\dot r^\alpha}, \partial_{p_i}\}$ is equivalent to

$$-\left(C^b_{fa}\frac{\partial l}{\partial \xi^b}\xi^f - \frac{d}{dt}\left(\frac{\partial l}{\partial \xi^a}\right)\right)A^a_\alpha dr^\alpha. \tag{5.9.51}$$

If we compute

$$
\begin{aligned}
&-\left(C^b_{fa}\frac{\partial l}{\partial \xi^b}\xi^f - \frac{d}{dt}\left(\frac{\partial l}{\partial \xi^a}\right)\right)A^a_\alpha dr^\alpha + \xi^a d\left(\frac{\partial l}{\partial \xi^a}\right) \\
&\quad + i_{(\dot r^\alpha \partial_{r^\alpha} + \ddot r^\alpha \partial_{\dot r^\alpha} + \dot p_i \partial p_i)}\left(dr^\alpha \wedge d\left(\frac{\partial l}{\partial \dot r^\alpha}\right)\right)
\end{aligned}
$$

and equate its terms with the corresponding terms of $d\overline{E}_{\overline{\mathcal{K}}}$ which is the same as $dE_{\mathcal{K}}$, we have the following equations after some computations

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -C^b_{da}\frac{\partial l}{\partial \xi^b}\xi^d A^a_\alpha - \frac{\partial l}{\partial \xi^a}\left(\dot{A}^a_\alpha - \frac{\partial A^a_\beta}{\partial r^\alpha}\dot{r}^\beta + \frac{\partial \Gamma^{ai} p_i}{\partial r^\alpha}\right).$$

After plugging in the constraint $\xi = -A\dot{r} + \Gamma p$ and simplify, we get the desired reduced equations

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b}(B^b_{\alpha\beta}\dot{r}^\beta + F^{bi}p_i), \tag{5.9.52}$$

where

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha \tag{5.9.53}$$

$$F^{bi}_\alpha = \frac{\partial \Gamma^{bi}}{\partial r^\alpha} - C^b_{ad}A^a_\alpha \Gamma^{di}. \tag{5.9.54}$$

In an orthogonal body frame where we choose our moving basis $e_b(g,r)$ to be orthogonal, that is, the corresponding generators $[e_b(g,r)]_Q$ are orthogonal in the given kinetic energy metric (actually, all that is needed is that the vectors in the set of basis vectors corresponding to the subspace $\mathcal{S}_q$ be orthogonal to the remaining basis vectors), the momentum equation (5.9.42) can be written as (see [BKMM])

$$\frac{d}{dt}p_i = C^j_{hi}I^{hl}p_j p_l + \mathcal{D}^j_{i\alpha}\dot{r}^\alpha p_j + \mathcal{D}_{\alpha\beta i}\dot{r}^\alpha \dot{r}^\beta, \tag{5.9.55}$$

where

$$\mathcal{D}^j_{i\alpha} = -C^j_{ai}A^a_\alpha + \gamma^j_{i\alpha} + \lambda_{a'\alpha}C^{a'}_{li}I^{lj} \tag{5.9.56}$$

$$\mathcal{D}_{\alpha\beta i} = \lambda_{a'\alpha}(-C^{a'}_{ai}A^a_\beta + \gamma^{a'}_{i\beta}). \tag{5.9.57}$$

Here $\gamma^c_{b\alpha}$ and $\lambda_{a'\alpha}$ are defined by

$$\frac{\partial e_b}{\partial r^\alpha} = \gamma^c_{b\alpha}e_c \tag{5.9.58}$$

$$\lambda_{a'\alpha} = \frac{\partial l}{\partial \xi^{a'}\partial \dot{r}^\alpha} - \frac{\partial l}{\partial \xi^{a'}\partial \xi^b}A^b_\alpha. \tag{5.9.59}$$

Notice that while the summation range of $a, b, c, d...$ are over all Lie algebra element (1 to k). those over $i, j, l, ...$ are the restricted (constrained) range (1 to m) and those over $a', b', ...$ run from $m+1$ to k (which correspond to the symmetry directions not aligned with the constraints).

Similarly we can rewrite the above reduced Lagrange-d'Alembert equations as in Theorem 5.7.3 of the text.

Remarks

1. A careful reading of the proof of Theorem 5.9.2 shows that the Hamiltonian reduction procedure still works as long as the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ is invertible. This is important because in some examples like the bicycle the Legendre transform $\mathbb{F}L$ is singular, but its restriction to the constraint submanifold $\mathcal{D}$ is invertible and the Hamiltonian reduction procedure is also applicable.

2. In many examples like the snakeboard and the bicycle, the constraints satisfy a special condition, namely, they involve only the velocities of the group variables $\dot{g}$ and are independent of the velocities of the shape variables $\dot{r}$ (see equations (5.9.9) and (5.9.10)). Under this special condition, the distribution $\mathcal{K}$ in equation (5.9.39) can be represented by
$$\mathcal{K} = \operatorname{span}\{g^a_d e^d_i \partial_{g^a}, \partial_r, \partial_{\dot{r}}, \partial_p\}. \tag{5.9.60}$$
This representation simplifies the computation for finding the reduced equations because the restriction of the one-form (5.9.48) to the subdistribution $\overline{\mathcal{K}}$ spanned by $\{\partial_r, \partial_{\dot{r}}, \partial_p\}$ will equal to zero. Hence in pushing down $i_{X_\mathcal{K}}\Omega_\mathcal{D}$ in (5.9.46) to $\overline{\mathcal{K}}$, we can simply omit the one-form (5.9.48). Below we will use this simplified procedure for the examples of the snakeboard and the bicycle.

3. Since the momentum equation is central to the theory of nonholonomic mechanical systems with symmetry, we make a few additional remarks about it. Before that, we state the following proposition:

5.9.3 Proposition. *For a nonholonomic mechanical system with symmetry, we have*
$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i}\right)(\xi^q_Q)^i = \frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi^q_Q)^i\right) - \frac{\partial L}{\partial \dot{q}^i}\left(\frac{d}{dt}\xi^q\right)^i_Q \tag{5.9.61}$$
where $\xi^q \in \mathfrak{g}^q$ *.*

Proof: Choose a section of $\mathfrak{g}^\mathcal{D}$ and apply the chain rule to give
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(\xi^q_Q)^i\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)(\xi^q_Q)^i + \frac{\partial L}{\partial \dot{q}^i}\left((T\xi^q_Q \cdot \dot{q})^i + \left(\frac{d}{dt}\xi^q\right)^i_Q\right).$$
Invariance of the Lagrangian implies that
$$L(exp(s\xi^q) \cdot q, \exp(s\xi^q) \cdot \dot{q}) = L(q, \dot{q}).$$
Differentiating this expression and evaluating it at $s = 0$, we get
$$\frac{\partial L}{\partial \dot{q}^i}(\xi^q_Q)^i + \frac{\partial L}{\partial \dot{q}^i}(T\xi^q_Q \cdot \dot{q})^i = 0.$$

After eliminating the term $\frac{\partial L}{\partial \dot{q}^i}(T\xi_Q^q \cdot \dot{q})^i$ from the above two equations, we arrive at the desired result. ■

The above equation can be rewritten as

$$\left\langle (dE - i_X\Omega_L)|_{\mathcal{D}}, (\xi_Q^q)' \right\rangle = \frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i} \right) (\xi_Q^q)^i \right) - \frac{\partial L}{\partial \dot{q}^i} \left(\frac{d}{dt}\xi^q \right)_Q^i, \quad (5.9.62)$$

where $(\xi_Q^q)' \in \mathcal{T} \cap \mathcal{K}$ and $T\tau_Q((\xi_Q^q)') = \xi_Q^q$. Since both the energy function E and the submanifold $\mathcal{D}$ are G-invariant, the left hand side of the above equation reduces to $\Omega_{\mathcal{D}}(X_{\mathcal{D}}, (\xi_Q^q)')$ and hence any vector field $X_{\mathcal{D}}$ which takes values in $\mathcal{W} = \mathcal{K} \cap (\mathcal{T} \cap \mathcal{K})^\perp$ will make the left hand side zero and hence must satisfy the momentum equation (5.9.20)

$$\frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}^i} \right) (\xi_Q^q)^i \right) - \frac{\partial L}{\partial \dot{q}^i} \left(\frac{d}{dt}\xi^q \right)_Q^i = 0, \quad (5.9.63)$$

as we have already seen in the proof of Theorem 5.9.2.

In showing that the vector field $X_{\mathcal{H}}$, which satisfies the equation

$$i_{X_{\mathcal{H}}}\Omega_{\mathcal{H}} = dH_{\mathcal{H}},$$

must lie in the subdistribution $\mathcal{U}$, one might think that any vector field $Y \in \mathcal{V} \cap \mathcal{H}$ can be expressed as a linear combination of infinitesimal generators (generated by fixed Lie algebra elements). But this is not the case, as we have pointed out earlier in the Lagrangian side, in general $(\xi_Q^q)'$ is the (vertical) lift of a section of the bundle $\mathcal{S}$ (generated by a section of the bundle $\mathfrak{g}^{\mathcal{D}}$). This is also true on the Hamiltonian side.

Example: The Snakeboard Revisited Now we return to the snakeboard and discuss the role of the symmetry group $G = SE(2)$. Recall from our earlier discussion that the Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mr^2\dot{\theta}^2 + + \frac{1}{2}J_0\dot{\psi}^2 + J_0\dot{\psi}\dot{\theta} + J_1\dot{\phi}_1^2, \quad (5.9.64)$$

which is independent of the configuration of the board and hence it is invariant to all possible group actions.

The Constraint Submanifold. The condition of rolling without slipping gives rise to the constraint one-forms

$$\begin{aligned} \omega_1(q) &= -\sin(\theta + \phi)dx + \cos(\theta + \phi)dy - r\cos\phi d\theta \\ \omega_2(q) &= -\sin(\theta - \phi)dx + \cos(\theta - \phi)dy + r\cos\phi d\theta, \end{aligned}$$

which are invariant under the $SE(2)$ action. The constraints determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \mathrm{span}\{\partial_\psi, \partial_\phi, a\partial_x + b\partial_y + c\partial_\theta\},$$

where $a = -2r\cos^2\phi\cos\theta, b = -2r\cos^2\phi\sin\theta, c = \sin 2\phi$. The tangent space to the orbits of the $SE(2)$ action is given by

$$T_q(\mathrm{Orb}(q)) = \mathrm{span}\{\partial_x, \partial_y, \partial_\theta\}.$$

The intersection between the tangent space to the group orbits and the constraint distribution is thus given by

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\mathrm{Orb}(q)) = \mathrm{span}\{a\partial_x + b\partial_y + c\partial_\theta\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\mathrm{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\mathrm{Orb}(q)$ is one-dimensional, the section can be chosen to be

$$\xi_Q^q = a\partial_x + b\partial_y + c\partial_\theta,$$

which is invariant under the action of $SE(2)$ on Q. The nonholonomic momentum is thus given by

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i \\ &= ma\dot{x} + mb\dot{y} + mr^2 c\dot{\theta} + J_0 c\dot{\psi}. \end{aligned}$$

The kinematic constraints plus the momentum are given by

$$\begin{aligned} 0 &= -\sin(\theta+\phi)\dot{x} + \cos(\theta+\phi)\dot{y} - r\cos\phi\dot{\theta} \\ 0 &= -\sin(\theta-\phi)\dot{x} + \cos(\theta-\phi)\dot{y} + r\cos\phi\dot{\theta} \\ p &= -2mr\cos^2\phi\cos\theta\dot{x} - 2mr\cos^2\phi\sin\theta\dot{y} \\ &\quad + mr^2\sin 2\phi\dot{\theta} + J_0\sin 2\phi\dot{\psi}. \end{aligned}$$

Adding, subtracting, and scaling these equations, we can write (away from the point $\phi = \pi/2$),

$$\begin{bmatrix} \cos\theta\dot{x} + \sin\theta\dot{y} \\ -\sin\theta\dot{x} + \cos\theta\dot{y} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} -\dfrac{J_0}{2mr}\sin 2\phi\dot{\psi} \\ 0 \\ \dfrac{J_0}{mr^2}\sin^2\phi\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{-1}{2mr}p \\ 0 \\ \dfrac{\tan\phi}{2mr^2}p \end{bmatrix}. \tag{5.9.65}$$

These equations have the form

$$g^{-1}\dot{g} + A(r)\dot{r} = \Gamma(r)p$$

where

$$\begin{aligned} A(r) &= -\frac{J_0}{2mr}\sin 2\phi e_x\, d\psi + \frac{J_0}{mr^2}\sin^2\phi e_\theta\, d\psi \\ \Gamma(r) &= \frac{-1}{2mr}e_x + \frac{1}{2mr^2}\tan\phi\, e_\theta. \end{aligned}$$

These are precisely the terms which appear in the nonholonomic connection relative to the (global) trivialization (r, g).

After applying the constrained Legendre transformation and its inverse to the constraint equations (5.9.65), we have

$$\begin{bmatrix} \cos\theta p_x + \sin\theta p_y \\ -\sin\theta p_x + \cos\theta p_y \\ p_\theta \end{bmatrix} + \begin{bmatrix} -\dfrac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)} p_\psi \\ 0 \\ -\dfrac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)} p_\psi \end{bmatrix} = \begin{bmatrix} \dfrac{-mr}{2(mr^2 - J_0\sin^2\phi)} p \\ 0 \\ \dfrac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)} p \end{bmatrix}, \tag{5.9.66}$$

where

$$p = -2r\cos^2\phi\cos\theta p_x - 2r\cos^2\phi\sin\theta p_y + \sin 2\phi p_\theta$$

and is SE(2)-invariant.

Therefore, the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$\begin{aligned} p_x &= \frac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)} p_\psi \cos\theta - \frac{mr}{2(mr^2 - J_0\sin^2\phi)} p\cos\theta \\ p_y &= \frac{mr\sin\phi\cos\phi}{(mr^2 - J_0\sin^2\phi)} p_\psi \sin\theta - \frac{mr}{2(mr^2 - J_0\sin^2\phi)} p\sin\theta \\ p_\theta &= \frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)} p_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)} p\,. \end{aligned}$$

It is a submanifold in T^*Q and we can use $(x, y, \theta, \psi, \phi, p_\psi, p_\phi, p)$ as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V} \cap \mathcal{H}$ and $\mathcal{U}$. With the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \operatorname{span}\{-2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\psi}, \partial_{p_\phi}, \partial_p\} \tag{5.9.67}$$

and the subdistribution $\mathcal{V} \cap \mathcal{H}$ is

$$\mathcal{V} \cap \mathcal{H} = \operatorname{span}\{-2r\cos^2\phi\cos\theta\partial_x - 2r\cos^2\phi\sin\theta\partial_y + \sin 2\phi\partial_\theta\}. \tag{5.9.68}$$

As for the subdistribution $\mathcal{U}$, we first calculate the two form $\Omega_{\mathcal{M}}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$
\begin{aligned}
\Omega_{\mathcal{M}} &= dx \wedge dp_x + dy \wedge dp_y + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\
&= (\cos\theta dx + \sin\theta dy) \wedge \left(\frac{mr \sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)} dp_\psi \right. \\
&\quad \left. - \frac{mr}{2(mr^2 - J_0 \sin^2\phi)} dp \right) \\
&\quad + (\cos\theta dx + \sin\theta dy) \wedge \left(\frac{mr(mr^2 \cos 2\phi + J_0 \sin^2\phi)}{(mr^2 - J_0 \sin^2\phi)^2} p_\psi d\phi \right. \\
&\quad \left. - \frac{mr J_0 \sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)^2} p d\phi \right) \\
&\quad + d\theta \wedge \left(\frac{mr^2 \cos^2\phi}{(mr^2 - J_0 \sin^2\phi)} dp_\psi + \frac{(mr^2 - J_0) \tan\phi}{2(mr^2 - J_0 \sin^2\phi)} dp \right) \\
&\quad + d\theta \wedge \left(\frac{mr^2 (J_0 - mr^2) \sin 2\phi}{(mr^2 - J_0 \sin^2\phi)^2} p_\psi d\phi \right. \\
&\quad \left. + \frac{(mr^2 - J_0)(mr^2 \sec^2\phi + J_0 \tan^2\phi \cos 2\phi)}{2(mr^2 - J_0 \sin^2\phi)^2} p d\phi \right) \\
&\quad + (-\sin\theta dx + \cos\theta dy) \wedge \left(\frac{mr \sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)} p_\psi \right. \\
&\quad \left. - \frac{mr}{2(mr^2 - J_0 \sin^2\phi)} p \right) d\theta \\
&\quad + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \,.
\end{aligned}
$$

Since $\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^\perp = \ker\{ i_{(\mathcal{V}\cap\mathcal{H})} \Omega_{\mathcal{H}} \}$, we need to calculate $i_{(\mathcal{V}\cap\mathcal{H})} \Omega_{\mathcal{M}}$, and restrict it to $\mathcal{H}$:

$$
\begin{aligned}
& i_{(\mathcal{V}\cap\mathcal{H})} \Omega_{\mathcal{H}} = \\
&\quad - 2r \cos^2\phi \left(\frac{mr \sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)} dp_\psi - \frac{mr}{2(mr^2 - J_0 \sin^2\phi)} dp \right) \\
&\quad - 2r \cos^2\phi \left(\frac{mr(mr^2 \cos 2\phi + J_0 \sin^2\phi)}{(mr^2 - J_0 \sin^2\phi)^2} p_\psi d\phi - \frac{mr J_0 \sin 2\phi}{2(mr^2 - J_0 \sin^2\phi)^2} p d\phi \right) \\
&\quad + \sin 2\phi \left(\frac{mr^2 \cos^2\phi}{(mr^2 - J_0 \sin^2\phi)} dp_\psi + \frac{(mr^2 - J_0) \tan\phi}{2(mr^2 - J_0 \sin^2\phi)} dp \right) \\
&\quad + \sin 2\phi \left(\frac{mr^2 (J_0 - mr^2) \sin 2\phi}{(mr^2 - J_0 \sin^2\phi)^2} p_\psi d\phi \right. \\
&\quad \left. + \frac{(mr^2 - J_0)(mr^2 \sec^2\phi + J_0 \tan^2\phi \cos 2\phi)}{2(mr^2 - J_0 \sin^2\phi)^2} p d\phi \right) \\
&\qquad = dp - \frac{2mr^2 \cos^2\phi}{mr^2 - J_0 \sin^2\phi} p_\psi d\phi + \frac{(mr^2 + J_0 \cos 2\phi) \tan\phi}{mr^2 - J_0 \sin^2\phi} p d\phi \,.
\end{aligned}
$$

Hence,

$$\mathcal{U} = \ker\left\{dp - \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}p_\psi d\phi + \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi}pd\phi\right\}. \tag{5.9.69}$$

The Reconstruction and Momentum Equations A vector field $X_\mathcal{U}$ taking values in $\mathcal{U}$ must be of the form

$$X_\mathcal{U} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p \tag{5.9.70}$$

where

$$\begin{aligned}\dot{x} &= \frac{J_0}{2mr}\sin 2\phi\dot{\psi}\cos\theta - \frac{1}{2mr}p\cos\theta\\ \dot{y} &= \frac{J_0}{2mr}\sin 2\phi\dot{\psi}\sin\theta - \frac{1}{2mr}p\sin\theta\\ \dot{\theta} &= -\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} + \frac{\tan\phi}{2mr^2}p\end{aligned}$$

and

$$\dot{p} = \frac{2mr^2\cos^2\phi}{mr^2 - J_0\sin^2\phi}p_\psi\dot{\phi} - \frac{(mr^2 + J_0\cos 2\phi)\tan\phi}{mr^2 - J_0\sin^2\phi}p\dot{\phi}\,. \tag{5.9.71}$$

The equations for $\dot{x}$, $\dot{y}$ and $\dot{\theta}$ are the same reconstruction equations as equations (5.9.65) and the last one for $\dot{p}$ is the momentum elution on the Hamiltonian side. As noted in [BKMM], the momentum p is the angular momentum of the system about the point P shown in figure 5.9.2.

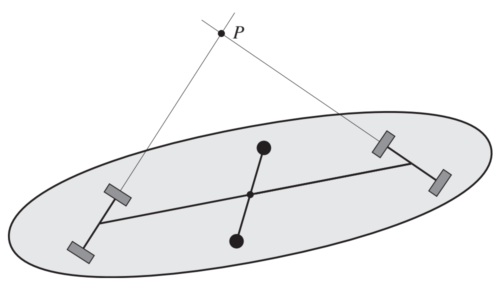

FIGURE 5.9.2. The momentum p is the angular momentum of the snakeboard system about the point P.

It can be checked that the momentum equation (5.9.71) is equivalent to the equation (5.9.11) via a change of variables with

$$\begin{aligned}p &= -2r\cos^2\phi\cos\theta p_x - 2r\cos^2\phi\sin\theta p_y + \sin 2\phi p_\theta\\ &= \frac{2(mr^2 - J_0\sin^2\phi)\cot\phi}{mr^2 - J_0}p_\theta - \frac{2mr^2\cos^2\phi\cot\phi}{mr^2 - J_0}p_\psi\end{aligned}$$

as the key link. Similarly the two full sets of equations of motion in both section 5.9 and this section are also related in the same way.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$i_{X_{\mathcal{H}}}\Omega_{\mathcal{M}} = dH_{\mathcal{M}}, \tag{5.9.72}$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $i_{X_{\mathcal{H}}}\Omega_{\mathcal{M}}$

$$\begin{aligned}
&i_{X_{\mathcal{H}}}\Omega_{\mathcal{M}} = \\
&(\dot{x}\cos\theta + \dot{y}\sin\theta)\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&+(\dot{x}\cos\theta + \dot{y}\sin\theta)\left(\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi\right. \\
&\left.-\frac{mrJ_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&+\dot{\theta}\left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&+\dot{\theta}\left(\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi\right. \\
&\left.+\frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&+\dot{\psi}dp_\psi + \dot{\phi}dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi \\
&-\dot{\theta}\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}p_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}p\right)(-\sin\theta dx + \cos\theta dy) \\
&-mr\left(\frac{mr^2\cos 2\phi + J_0\sin^2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi\dot{\phi}\right. \\
&\left.-\frac{J_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}p\dot{\phi}\right)(\cos\theta dx + \sin\theta dy) \\
&-\left(\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi\dot{\phi}\right. \\
&\left.+\frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}p\dot{\phi}\right)d\theta \\
&-\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}\dot{p}_\psi(\cos\theta dx + \sin\theta dy) - \frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}\dot{p}_\psi d\theta \\
&+\frac{mr}{2(mr^2 - J_0\sin^2\phi)}(\cos\theta dx + \sin\theta dy) - \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}\dot{p}d\theta.
\end{aligned}$$

As for $dH_{\mathcal{H}}$, recall that the constrained Hamiltonian $H_{\mathcal{M}}$ is

$$H_{\mathcal{M}} = \frac{mr^2}{2(mr^2-J_0)^2}\cot^2\phi(p_\theta - p_\psi)^2 + \frac{1}{2J_0}p_\psi^2 + \frac{1}{2(mr^2-J_0)}(p_\theta - p_\psi)^2 + \frac{1}{4J_1}p_\phi^2.$$

Notice that $H_{\mathcal{M}}$ is $SE(2)$-invariant and hence $H_{\mathcal{M}} = h_{\overline{\mathcal{M}}}$ where

$$\begin{aligned} h_{\overline{\mathcal{M}}} = &\frac{mr^2}{2}\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}p - \frac{\sin^2\phi}{2(mr^2-J_0\sin^2\phi)}p_\psi\right)^2 + \frac{1}{2J_0}p_\psi^2 \\ &+ \frac{mr^2-J_0}{2}\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}p - \frac{\sin^2\phi}{mr^2-J_0\sin^2\phi}p_\psi\right)^2 + \frac{1}{4J_1}p_\phi^2. \end{aligned}$$

Computing $dH_{\mathcal{M}} = dh_{\overline{\mathcal{M}}}$, we get

$$\begin{aligned} dh_{\overline{\mathcal{M}}} = &\frac{mr^2(p-\sin 2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)} \\ &\quad\times\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}dp - \frac{\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}dp_\psi\right) \\ &+ \frac{mr^2(p-\sin 2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)} \\ &\quad\times\left(pd\left(\frac{1}{2(mr^2-J_0\sin^2\phi)}\right) - p_\psi d\left(\frac{\sin 2\phi}{2(mr^2-J_0\sin^2\phi)}\right)\right) \\ &+ \frac{(mr^2-J_0)(\tan\phi p - 2\sin^2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)} \\ &\quad\times\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}dp - \frac{\sin^2\phi}{(mr^2-J_0\sin^2\phi)}dp_\psi\right) \\ &+ \frac{(mr^2-J_0)(\tan\phi p - 2\sin^2\phi p_\psi)}{2(mr^2-J_0\sin^2\phi)} \\ &\quad\times\left(pd\left(\frac{\tan\phi}{2(mr^2-J_0\sin^2\phi)}\right) - p_\psi d\left(\frac{\sin^2\phi}{(mr^2-J_0\sin^2\phi)}\right)\right) \\ &+ \frac{1}{J_0}p_\psi dp_\psi + + \frac{1}{2J_1}p_\phi dp_\phi. \end{aligned}$$

It is easy to check that $i_{X_{\mathcal{H}}}\Omega_{\mathcal{M}} = dH_{\mathcal{M}}$ is SE(2)-invariant, and vanishes on $\mathcal{V}\cap\mathcal{H}$ when restricted to $\mathcal{U}$. Hence both sides push down to $\overline{\mathcal{H}}$. The push down of $i_{X_{\mathcal{H}}}\Omega_{\mathcal{M}}$ is given by

$$\begin{aligned}
i_{X_{\overline{\mathcal{H}}}}\Omega_{\overline{\mathcal{H}}} = &\left(\frac{J_0}{2mr}\sin(2\phi)\dot{\psi} - \frac{1}{2mr}p\right) \\
&\times\left(\frac{mr\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)}dp_\psi - \frac{mr}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&+\left(\frac{J_0}{2mr}\sin(2\phi)\dot{\psi} - \frac{1}{2mr}p\right) \\
&\times\left(\frac{mr(mr^2\cos 2\phi + J_0\sin^2\phi)}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi - \frac{mrJ_0\sin 2\phi}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi\right) \\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right) \\
&\times\left(\frac{mr^2\cos^2\phi}{(mr^2 - J_0\sin^2\phi)}dp_\psi + \frac{(mr^2 - J_0)\tan\phi}{2(mr^2 - J_0\sin^2\phi)}dp\right) \\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right) \\
&\times\frac{mr^2(J_0 - mr^2)\sin 2\phi}{(mr^2 - J_0\sin^2\phi)^2}p_\psi d\phi \\
&+\left(\frac{-J_0}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right) \\
&\times\frac{(mr^2 - J_0)(mr^2\sec^2\phi + J_0\tan^2\phi\cos 2\phi)}{2(mr^2 - J_0\sin^2\phi)^2}pd\phi \\
&+\dot{\psi}dp_\psi + \dot{\phi}dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi.
\end{aligned}$$

Equating the terms of $dh_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{M}}}$ with those of the push down of $i_{X_{\mathcal{H}}}\Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$\dot{\psi} = -\frac{\tan\phi}{2(mr^2 - J_0\sin^2\phi)}p + \frac{mr^2}{J_0(mr^2 - J_0\sin^2\phi)}p_\psi \tag{5.9.73}$$

$$\dot{\phi} = \frac{p_\phi}{2J_1} \tag{5.9.74}$$

$$\dot{p}_\psi = 0 \tag{5.9.75}$$

$$\dot{p}_\phi = 0. \tag{5.9.76}$$

Notice that both the momentum equation (5.9.71) and the above set of reduced equations are independent of the group elements of the symmetry group $SE(2)$. If we add in the set of reconstruction equations (5.9.65), we recover the full dynamics of the system, and in a form that is suitable for control theoretical purposes.

Finding the Reduced Equations on the Lagrangian Side We can derive the reduced Lagrange-d'Alembert equations as follows: Here we will use the equations (5.9.52).

$$\frac{d}{dt}\left(\frac{\partial l_c}{\partial \dot{r}^\alpha}\right) - \frac{\partial l_c}{\partial r^\alpha} = -\frac{\partial l}{\partial \xi^b}(B^b_{\alpha\beta}\dot{r}^\beta + F^{bi}p_i), \tag{5.9.77}$$

where

$$B^b_{\alpha\beta} = \frac{\partial A^b_\alpha}{\partial r^\beta} - \frac{\partial A^b_\beta}{\partial r^\alpha} - C^b_{ac}A^a_\beta A^c_\alpha \quad \text{and} \quad F^{bi}_\alpha = \frac{\partial \Gamma^{bi}}{\partial r^\alpha} - C^b_{ad}A^a_\alpha \Gamma^{di}.$$

From the Lagrangian L, we find the reduced Lagrangian

$$l(r,\dot{r},\xi) = \frac{1}{2}m((\xi^1)^2 + (\xi^2)^2) + \frac{1}{2}mr^2(\xi^3)^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_0\dot{\psi}(\xi^3) + J_1\dot{\phi}^2, \tag{5.9.78}$$

where $\xi = g^{-1}\dot{g}$. After plugging in the constraints (5.9.65), we have the constrained reduced Lagrangian

$$l_c(r,\dot{r},p) = -\frac{J_0^2}{2mr^2}\sin^2\phi\dot{\psi}^2 + \frac{1}{8mr^2}\sec^2\phi p^2 + \frac{1}{2}J_0\dot{\psi}^2 + +J_1\dot{\phi}^2. \tag{5.9.79}$$

Let us find all the ingredients of the above equations:

$$\begin{aligned}
\frac{\partial l}{\partial \xi^1} &= m\xi^1 = m\left(\frac{J_0}{2mr}\sin 2\phi\dot{\psi} - \frac{1}{2mr}p\right)\\
\frac{\partial l}{\partial \xi^2} &= m\xi^2 = 0\\
\frac{\partial l}{\partial \xi^3} &= mr^2\left(-\frac{J_0}{mr^2}\sin^2\phi\dot{\psi} + \frac{\tan\phi}{2mr^2}p\right) + J_0\dot{\psi};
\end{aligned}$$

since $\frac{\partial l}{\partial \xi^2} = 0$, we do not need to compute $B^2_{\alpha\beta}$ and F^2_α (notice that $i = 1$). Also it is straightforward to find

$$\begin{aligned}
B^1_{12} &= \frac{\partial}{\partial\phi}\left(-\frac{J_0}{2mr}\sin 2\phi\right) = -\frac{J_0}{mr}\cos 2\phi\\
B^3_{12} &= \frac{\partial}{\partial\phi}\left(\frac{J_0}{mr}\sin^2\phi\right) = \frac{J_0}{mr}\sin 2\phi\\
F^3_2 &= \frac{\partial}{\partial\phi}\left(\frac{\tan\phi}{2mr^2}\right) = \frac{\sec^2\phi}{2mr^2},
\end{aligned}$$

and $F^1_1 = F^3_1 = F^1_2 = 0$. Substituting into (5.9.77), we get the reduced equations after some computations

$$\left(1 - \frac{J_0}{mr^2}\sin^2\phi\right)\ddot{\psi} = \frac{J_0}{2mr^2}\sin 2\phi\dot{\psi}\dot{\phi} - \frac{J_0}{2mr^2}\dot{\phi}p \tag{5.9.80}$$

$$J_1\ddot{\phi} = 0\,. \tag{5.9.81}$$

It is easy to check that these two equations are equivalent to the set of reduced equations (5.9.73)-(5.9.76) on the Hamiltonian side through the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$.

Example: The Bicycle. Control of the bicycle is a rich problem offering a number of considerable challenges of current research interest in the area of mechanical and robotic control. The bicycle is an underactuated system, subject to nonholonomic contact constraints associated with the rolling constraints on the front and rear wheels. It is unstable (except under certain combinations of fork geometry and speed) when not controlled. It is also, when considered to traverse flat ground, a system subject to symmetries; its Lagrangian and constraints are invariant with respect to translations and rotations in the ground plane.

Here a simplified bicycle model will be considered. The wheels of the bicycle are considered to have negligible inertia moments, mass, radii, and width, and roll without side or longitudinal slip. The vehicle is assumed to have a fixed steering axis that is perpendicular to the flat ground when the bicycle is upright. For simplicity we concern ourselves with a point mass bicycle. The rigid frame of the bicycle will be assumed to be symmetric about a plane containing the rear wheel.

Consider a ground fixed inertial reference frame with x and y axis in the ground plane and z-axis perpendicular to the ground plane in the direction opposite to gravity. The intersection of the vehicle's plane of symmetry with the ground plane forms a contact line. The contact line is rotated about the z-direction by a yaw angle θ. The contact line is considered directed, with its positive direction from the rear to the front of the vehicle. The yaw angle θ is zero when the contact line is in the x-direction. The angle that the bicycle's plane of symmetry makes with the vertical direction is the roll angle $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Front and rear wheel contacts are constrained to have velocities parallel to the lines of intersection of their respective wheel planes and the ground plane, but free to turn about an axis through the wheel/ground contact and parallel to the z-axis. Let $\sigma \in (-\frac{\pi}{2}, \frac{\pi}{2})$ be the steering angle between the front wheel plane/ground plane intersection and the contact line. With σ we associate a moment of inertia J which depends both on ψ and σ. We will parameterize the steering angle by $\phi := \tan\sigma/b$. For more details, see Getz and Marsden [1995]. See figure 5.9.3.

The configuration space is $Q = SE(2) \times S^1 \times S^1$ and the Lagrangian $L : TQ \to \mathbb{R}$ is the total kinetic energy minus potential energy of the system and is given by

$$\begin{aligned} L = {} & -mga\cos\psi + \frac{1}{2}J(\psi,\phi)\dot{\phi}^2 \\ & + \frac{m}{2}\Big((\cos\theta\dot{x} + \sin\theta\dot{y} + a\sin\psi\dot{\theta})^2 \\ & + (-\sin\theta\dot{x} + \cos\theta\dot{y} - a\cos\psi\dot{\psi} + c\dot{\theta})^2 + (-a\sin\psi\dot{\psi})^2\Big) \end{aligned}$$

where m is the mass of the bicycle, considered for simplicity to be a point mass, and $J(\psi,\phi)$ is the moment of inertia associated with the steering action. The nonholonomic constraints associated with the front and rear

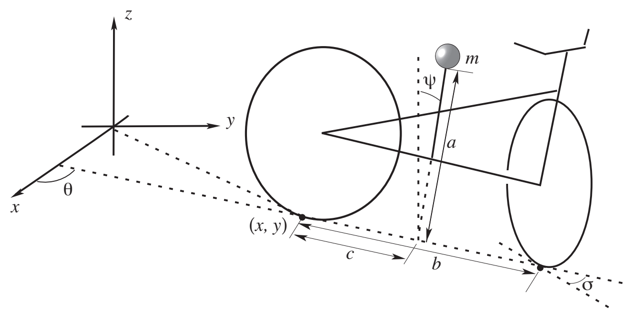

FIGURE 5.9.3. Notation for the bike.

wheels, assumed to roll without slipping, are expressed by

$$\begin{aligned} \dot{\theta} - \phi(\cos\theta\dot{x} + \sin\theta\dot{y}) &= 0 \\ -\sin\theta\dot{x} + \cos\theta\dot{y} &= 0. \end{aligned}$$

Clearly both the Lagrangian and the constraints are invariant under the $SE(2)$ action.

Notice that the Legendre transform $\mathbb{F}L$ is singular but by remark 1 following Theorem 5.9.2 the Hamiltonian procedure still works because the constrained Legendre transform $\mathbb{F}L|\mathcal{D}$ is invertible.

The Constraint Submanifold. The constraints above give rise to the constraint one-forms

$$\begin{aligned} \omega_1(q) &= d\theta - \phi\cos\theta dx - \phi\sin\theta dy \\ \omega_2(q) &= -\sin\theta dx + \cos\theta dy \end{aligned}$$

which determine the kinematic distribution $\mathcal{D}_q$:

$$\mathcal{D}_q = \operatorname{span}\{\partial_\psi, \partial_\phi, \cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}.$$

The tangent space to the orbits of the $SE(2)$ action is given by

$$T_q(\operatorname{Orb}(q)) = \operatorname{span}\{\partial_x, \partial_y, \partial_\theta\},$$

and the intersection between the tangent space to the group orbits and the constraint distribution is thus given by

$$\mathcal{S}_q = \mathcal{D}_q \cap T_q(\operatorname{Orb}(q)) = \operatorname{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}.$$

The momentum can be constructed by choosing a section of $\mathcal{S} = \mathcal{D} \cap T\operatorname{Orb}$ regarded as a bundle over Q. Since $\mathcal{D}_q \cap T_q\operatorname{Orb}(q)$ is one-dimensional, the section can be chosen to be

$$\xi_Q^q = \cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta,$$

which is invariant under the action of $SE(2)$ on Q. The nonholonomic momentum map is thus given by

$$\begin{aligned}
p &= \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i \\
&= m(\dot{x} + a\sin\psi\cos\theta\dot{\theta} + a\cos\psi\sin\theta\dot{\psi} - c\sin\theta\dot{\theta})\cos\theta \\
&\quad + m(\dot{y} + a\sin\psi\sin\theta\dot{\theta} - a\cos\psi\cos\theta\dot{\psi} + c\cos\theta\dot{\theta})\sin\theta \\
&\quad + m(\cos\theta\dot{x} + \sin\theta\dot{y} + a\sin\psi\dot{\theta})a\phi\sin\psi \\
&\quad + m(-\sin\theta\dot{x} + \cos\theta\dot{y} - a\cos\psi\dot{\psi} + c\dot{\theta})c\phi.
\end{aligned}$$

The kinematic constraints plus the momentum are given by

$$\begin{aligned}
0 &= \xi^3 - \phi\xi^1 \\
0 &= \xi^2 \\
p &= m(\xi^1 + a\sin\psi\xi^3) + ma\phi\sin\psi(\xi^1 + a\sin\psi\xi^3) \\
&\qquad m\phi(c\xi^2 - ca\cos\psi\dot{\psi} + c^2\xi^3)
\end{aligned}$$

where

$$\begin{aligned}
\xi^1 &= \cos\theta\dot{x} + \sin\theta\dot{y} \\
\xi^2 &= -\sin\theta\dot{x} + \cos\theta\dot{y} \\
\xi^3 &= \dot{\theta}
\end{aligned}$$

Adding, subtracting, and scaling these equations, we can write

$$\begin{bmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{bmatrix} + \begin{bmatrix} -\dfrac{ca\phi\cos\psi}{K}\dot{\psi} \\ 0 \\ -\dfrac{ca\phi^2\cos\psi}{K}\dot{\psi} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{mK}p \\ 0 \\ \dfrac{\phi}{mK}p \end{bmatrix} \tag{5.9.82}$$

where

$$K = (1 + a\phi\sin\psi)^2 + c^2\phi^2. \tag{5.9.83}$$

These equations have the form

$$g^{-1}\dot{g} + A(r)\dot{r} = \Gamma(r)p.$$

Next find the Legendre transform $\mathbb{F}L$ and restrict it to the constraint submanifold $\mathcal{D} \subset TQ$, we get

$$\begin{aligned}
p_x &= m(1 + a\phi\sin\psi)\xi^1\cos\theta - m(c\phi\xi^1 - a\cos\psi\dot{\psi})\sin\theta \\
p_y &= m(1 + a\phi\sin\psi)\xi^1\sin\theta + m(c\phi\xi^1 - a\cos\psi\dot{\psi})\cos\theta \\
p_\theta &= ma\sin\psi(1 + a\phi\sin\psi)\xi^1 + m(c^2\phi\xi^1 - ca\cos\psi\dot{\psi}) \\
p_\psi &= ma^2\dot{\psi} - mac\cos\psi\phi\xi^1 \\
p_\phi &= J(\psi,\phi)\dot{\phi}.
\end{aligned}$$

After applying the constrained Legendre transformation $\mathbb{F}L|\mathcal{D}$ and its inverse to the constraint equations (5.9.82), we have

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} -\dfrac{c\phi\cos\psi(1+a\phi\sin\psi)}{F}\dfrac{p_\psi}{a} \\ \dfrac{(1+a\phi\sin\psi)^2\cos\psi}{F}\dfrac{p_\psi}{a} \\ \dfrac{c\cos\psi(1+a\phi\sin\psi)}{F}\dfrac{p_\psi}{a} \end{bmatrix} = \begin{bmatrix} \dfrac{1+a\phi\sin\psi}{F}p \\ \dfrac{c\phi\sin^2\psi}{F}p \\ \dfrac{(1+a\phi\sin\psi)a\sin\psi + c^2\phi\sin^2\psi}{F}p \end{bmatrix}, \tag{5.9.84}$$

where

$$\begin{aligned} \mu_1 &= \cos\theta p_x + \sin\theta p_y \\ \mu_2 &= -\sin\theta p_x + \cos\theta p_y \\ \mu_3 &= p_\theta \end{aligned}$$

and

$$F = (1+a\phi\sin\psi)^2 + c^2\phi^2\sin^2\psi \tag{5.9.85}$$
$$p = p_x\cos\theta + p_y\sin\theta + p_\theta\phi. \tag{5.9.86}$$

Therefore, the constraint submanifold $\mathcal{M} \subset T^*Q$ is defined by

$$\begin{aligned} p_x &= \mu_1\cos\theta - \mu_2\sin\theta \\ p_y &= \mu_1\sin\theta + \mu_2\cos\theta \\ p_\theta &= \mu_3. \end{aligned}$$

It is a submanifold in T^*Q and we can use $(x, y, \theta, \psi, \phi, p_\psi, p_\phi, p)$ as its induced local coordinates.

The Distributions $\mathcal{H}, \mathcal{V}\cap\mathcal{H}$ and $\mathcal{U}$. Using the induced coordinates, the distribution $\mathcal{H}$ on $\mathcal{M}$ is

$$\mathcal{H} = \operatorname{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta, \partial_\psi, \partial_\phi, \partial_{p_\psi}, \partial_{p_\phi}, \partial_p\} \tag{5.9.87}$$

and the subdistribution $\mathcal{V}\cap\mathcal{H}$ is

$$\mathcal{V}\cap\mathcal{H} = \operatorname{span}\{\cos\theta\partial_x + \sin\theta\partial_y + \phi\partial_\theta\}. \tag{5.9.88}$$

Notice that in the case of the bicycle, the constraints are independent of the velocities of the shape variables and hence the simplified procedure employed in the snakeboard is also used here.

As for the subdistribution $\mathcal{U}$, we first calculate the two-form $\Omega_{\mathcal{M}}$. After pulling back the canonical two-form of T^*Q to $\mathcal{M}$, we have

$$\begin{aligned}\Omega_{\mathcal{M}} &= dx \wedge dp_x + dy \wedge dp_y + d\theta \wedge dp_\theta + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi \\ &= (\cos\theta dx + \sin\theta dy) \wedge d\mu_1 + \mu_1(-\sin\theta dx + \cos\theta dy) \wedge d\theta \\ &\quad +(-\sin\theta dx + \cos\theta dy) \wedge d\mu_2 - \mu_2(\cos\theta dx + \sin\theta dy) \wedge d\theta \\ &\quad +d\theta \wedge d\mu_3 + d\psi \wedge dp_\psi + d\phi \wedge dp_\phi\,.\end{aligned}$$

Since $\mathcal{U} = (\mathcal{V} \cap \mathcal{H})^\perp = \ker\{i_{(\mathcal{V}\cap\mathcal{H})}\Omega_{\mathcal{H}}\}$, we need to calculate $i_{(\mathcal{V}\cap\mathcal{H})}\Omega_{\mathcal{M}}$, and restrict it to $\mathcal{H}$:

$$\begin{aligned}i_{(\mathcal{V}\cap\mathcal{H})}\Omega_{\mathcal{H}} &= d\mu_1 - \mu_1\phi(-\sin\theta dx + \cos\theta dy) \\ &\quad -\mu_2 d\theta + \mu_2\phi(\cos\theta dx + \sin\theta dy) + \phi d\mu_3 \\ &= d\mu_1 + \phi d\mu_3 \\ &= dp + \frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}d\phi \\ &\quad -\frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}p d\phi.\end{aligned}$$

Hence,

$$\begin{aligned}\mathcal{U} = \ker\{dp &+ \frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}d\phi \\ &- \frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}p d\phi\}.\end{aligned} \tag{5.9.89}$$

The Reconstruction and Momentum Equations. A vector field $X_{\mathcal{U}}$ taking values in $\mathcal{U}$ must be of the form

$$X_{\mathcal{U}} = \dot{x}\partial_x + \dot{y}\partial_y + \dot{\theta}\partial_\theta + \dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi\partial_{p_\psi} + \dot{p}_\phi\partial_{p_\phi} + \dot{p}\partial_p \tag{5.9.90}$$

where

$$\begin{aligned}\dot{x} &= \xi^1\cos\theta - \xi^2\sin\theta = \left(\frac{ca\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p\right)\cos\theta \\ \dot{y} &= \xi^1\sin\theta + \xi^2\cos\theta = \left(\frac{ca\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p\right)\sin\theta \\ \dot{\theta} &= \phi\xi^1 = \left(\frac{ca\phi^2\cos\psi}{K}\dot{\psi} + \frac{\phi}{mK}p\right)\end{aligned}$$

and

$$\dot{p} = -\frac{c\cos\psi(1 + a\phi\sin\psi)}{F}\frac{p_\psi}{a}\dot{\phi} + \frac{a\sin\psi(1 + a\phi\sin\psi) + c^2\phi\sin^2\psi}{F}p\dot{\phi}. \tag{5.9.91}$$

The $\dot{p}$ equation is the momentum equation on the Hamiltonian side. Similar to the example of the snakeboard, the momentum p equals the angular momentum of the system about a fixed point P that can be determined in the same way as in the case of the snakeboard. Notice also that the last equation can be written simply as $\dot{p} = \mu_3 \dot{\phi}$.

The Reduced Hamilton Equations. To find the remaining reduced equations, we need to compute

$$i_{X_{\mathcal{H}}} \Omega_{\mathcal{M}} = dH_{\mathcal{M}}, \tag{5.9.92}$$

restrict it to the subdistribution $\mathcal{U}$ and then push it down to the reduced constraint submanifold $\overline{\mathcal{M}}$. Let us first compute $i_{X_{\mathcal{H}}} \Omega_{\mathcal{M}}$

$$\begin{aligned}
&i_{X_{\mathcal{H}}} \Omega_{\mathcal{M}} = \\
&(\cos\theta \dot{x} + \sin\theta \dot{y}) d\mu_1 + \mu_1(-\sin\theta \dot{x} + \cos\theta \dot{y}) d\theta - \mu_1 \dot{\theta}(-\sin\theta dx + \cos\theta dy) \\
&+(-\sin\theta \dot{x} + \cos\theta \dot{y}) d\mu_2 - \mu_2(\cos\theta \dot{x} + \sin\theta \dot{y}) d\theta + \mu_2 \dot{\theta}(\cos\theta dx + \sin\theta dy) \\
&+\dot{\theta} d\mu_3 + \dot{\psi} dp_\psi + \dot{\phi} dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi \\
&-(i_{(\dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi \partial_{p_\psi} + \dot{p}_\phi \partial_{p_\phi} + \dot{p}\partial_p)} d\mu_1)(\cos\theta dx + \sin\theta dy) \\
&-(i_{(\dot{\psi}\partial_\psi + \dot{\phi}\partial_\phi + \dot{p}_\psi \partial_{p_\psi} + \dot{p}_\phi \partial_{p_\phi} + \dot{p}\partial_p)} d\mu_2)(-\sin\theta dx + \cos\theta dy).
\end{aligned}$$

As for $dH_{\mathcal{H}}$, we can find the constrained Hamiltonian $H_{\mathcal{M}}$ via the constrained Legendre transform and have

$$\begin{aligned}
H_{\mathcal{M}} =& mga\cos\psi + \frac{1}{2J} p_\phi^2 \\
&+ \frac{1}{2m}\left(\mu_1^2 + \mu_2^2 + \left(\frac{K\sin\psi}{F}\frac{p_\psi}{a} + \frac{c\phi \sin\psi\cos\psi}{F} p\right)^2\right).
\end{aligned}$$

Notice that $H_{\mathcal{M}}$ is $SE(2)$-invariant and hence $H_{\mathcal{M}} = h_{\overline{\mathcal{M}}}$. Compute $dH_{\mathcal{M}} = dh_{\overline{\mathcal{M}}}$ and we have

$$\begin{aligned}
&dh_{\overline{\mathcal{M}}} = \\
&-mga\sin\psi d\psi + \frac{1}{J} p_\phi dp_\phi - \frac{1}{2J^2} p_\phi^2 \left(\frac{\partial J}{\partial\psi} d\psi + \frac{\partial J}{\partial\phi} d\phi\right) \\
&+\frac{1}{m}\left(\mu_1 d\mu_1 + \mu_2 d\mu_2\right. \\
&+\left.\left(\frac{K\sin\psi}{F}\frac{p_\psi}{a} + \frac{c\phi\sin\psi\cos\psi}{F} p\right) d\left(\frac{K\sin\psi}{F}\frac{p_\psi}{a} + \frac{c\phi\sin\psi\cos\psi}{F} p\right)\right).
\end{aligned}$$

It can be checked that $i_{X_{\mathcal{H}}} \Omega_{\mathcal{M}} = dH_{\mathcal{M}}$ is $SE(2)$-invariant, and vanishes on $\mathcal{V} \cap \mathcal{H}$ when restricted to $\mathcal{U}$. Hence both sides push down to $\overline{\mathcal{H}}$. The push down of $i_{X_{\mathcal{H}}} \Omega_{\mathcal{M}}$ is given by

$$\begin{aligned}
i_{X_{\overline{\mathcal{H}}}} \Omega_{\overline{\mathcal{H}}} &= (\cos\theta \dot{x} + \sin\theta \dot{y}) d\mu_1 + \dot{\theta} d\mu_3 + \dot{\psi} dp_\psi + \dot{\phi} dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi \\
&= \xi^1 d\mu_1 + \xi^3 d\mu_3 + \dot{\psi} dp_\psi + \dot{\phi} dp_\phi - \dot{p}_\psi d\psi - \dot{p}_\phi d\phi.
\end{aligned}$$

Equating the terms of $dh_{\overline{\mathcal{H}}} = dh_{\overline{\mathcal{M}}}$ with those of the push down of $i_{X_{\mathcal{H}}}\Omega_{\mathcal{M}}$ gives the remaining reduced Hamilton equations:

$$\dot{\psi} = \frac{1}{ma}\left(\frac{K}{F}\frac{p_\psi}{a} + \frac{c\phi\cos\psi}{F}p\right) \tag{5.9.93}$$

$$\dot{\phi} = \frac{p_\phi}{J} \tag{5.9.94}$$

$$\begin{aligned}\dot{p}_\psi &= mga\sin\psi + \frac{1}{2J^2}p_\phi^2\frac{\partial J}{\partial\psi} + m(1 + a\phi\sin\psi)a\phi\cos\psi(\xi^1)^2 \\ &\quad + mca\phi\sin\psi\xi^1\dot{\psi}\end{aligned} \tag{5.9.95}$$

$$\dot{p}_\phi = \frac{1}{2J^2}\frac{\partial J}{\partial\phi}p_\phi^2, \tag{5.9.96}$$

where

$$\xi^1 = \frac{c\phi\cos\psi}{K}\dot{\psi} + \frac{1}{mK}p = \frac{c\phi\cos\psi}{mF}\frac{p_\psi}{a} + \frac{1}{mF}p$$

as defined earlier in (5.9.82). The first two equations are nothing but the inverse of the constrained Legendre transform. Notice that both the momentum equation (5.9.91) and the above set of reduced equations are independent of the group elements of the symmetry group $SE(2)$. If we add in the set of reconstruction equations (5.9.82), we recover the full dynamics of the system, and in a form that is suitable for control theoretical purposes. Methods developed in Koon and Marsden [1997a] and will be used to study the optimal control of the bicycle whose equations of motion have been found in this section.

6

Control of Mechanical and Nonholonomic Systems

6.3 Stabilization of the Brockett Canonical Form in the Case so(n)

Here we illustrate the stabilization algorithm of §6.3 of the text. We consider the so(n) systems (6.1.2—6.1.3). Let $\mathfrak{g} = so(n+1)$. Instead of working with $(n+1) \times (n+1)$ matrices, we identify $\mathfrak{h}$ with so(n) and we identify the subspace $\mathfrak{m}$ with $\mathbb{R}^n$. Since $\mathfrak{g}$ is of compact type, $\epsilon = -1$. For $x \in \mathbb{R}^n$, the operator $M(x) : \mathrm{so}(n) \to \mathrm{so}(n)$ is given by

$$M(x)Y = -[x, [x, Y]] = xx^TY + Yxx^T. \tag{6.3.50}$$

This satisfies a minimal polynomial equation:

$$M(x)^2Y = (x^Tx)M(x)Y. \tag{6.3.51}$$

For $Y \in \mathrm{so}(n)$, the operator $N(Y) : \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$N(Y)x = Y^TYx. \tag{6.3.52}$$

The control (6.3.11) of the text in this setting is given by

$$u = -\alpha x + \beta Yx + \gamma Y^TYx, \tag{6.3.53}$$

for $x \in \mathbb{R}^n$, $Y \in \mathrm{so}(n)$, and thus the system (6.3.12–6.3.13) in the text is

$$\dot{x} = -\alpha x + \beta Yx + \gamma Y^TYx \tag{6.3.54}$$

$$\dot{Y} = -\beta(xx^TY + Yxx^T) + \gamma(Y^TYxx^T - xx^TY^TY). \tag{6.3.55}$$

In Case III, with $\alpha = \gamma = 0$, we know from equation (6.3.32) of the text that the matrix $M(x)Y - (x^T x)Y$ evolves with constant spectrum. Since $x^T x$ remains constant throughout this case, this means that

$$Y - \frac{1}{x^T x} M(x)Y \tag{6.3.56}$$

has constant spectrum.

Assume now that the system is in its initial configuration for Step 3 of the algorithm in §6.3. Then $Y^T Y x = \lambda_* x$ where λ_* is the largest eigenvalue of $N(Y) = Y^T Y$. Observe that

$$\begin{aligned} M(x)Y = xx^T Y + Yxx^T &= xx^T Y + \frac{1}{\lambda_*} Yxx^T Y^T Y \\ &= \left(xx^T + \frac{1}{\lambda_*} Yxx^T Y^T \right) Y. \end{aligned}$$

We claim the symmetric operator $\frac{1}{x^T x}(xx^T + \frac{1}{\lambda_*} Yxx^T Y^T)$ is actually an orthogonal projector onto the subspace of $\mathbb{R}^n$ spanned by x and Yx. Indeed, using the antisymmetry of Y, we compute

$$\begin{aligned} \left(xx^T + \frac{1}{\lambda_*} Yxx^T Y^T \right)^2 &= (x^T x)xx^T + \frac{1}{{\lambda_*}^2} Yxx^T Y^T Yxx^T Y^T \\ &= (x^T x)(xx^T + \frac{1}{\lambda_*} Yxx^T Y^T). \end{aligned}$$

This establishes our claim. It follows that the operator given by (6.3.56) is the composition of the orthogonal projector onto the orthogonal complement of the subspace spanned by x and Yx with Y.

In particular, suppose v is an eigenvector of $Y^T Y$ corresponding to a nonzero eigenvalue $\lambda \neq \lambda_*$. Then Yx is also such an eigenvector, and $Yv \pm \lambda^{1/2} iv$ are complex eigenvectors of Y itself corresponding to the eigenvalues $\pm\lambda^{1/2} i$, respectively. Both v and Yv are orthogonal to the subspace spanned by x and Yx. The preceding discussion shows that

$$\begin{aligned} \left(Y - \frac{1}{x^T x} M(x)Y \right) v &= Yv \\ \left(Y - \frac{1}{x^T x} M(x)Y \right) Yv &= Y^2 v = -\lambda v. \end{aligned}$$

It follows that

$$\left(Y - \frac{1}{x^T x} M(x)Y \right) (Yv \pm \lambda^{1/2} iv) = \pm\lambda^{1/2} i (Yv \pm \lambda^{1/2} iv),$$

and thus $\pm\lambda^{1/2} i$ are eigenvalues of

$$Y - \frac{1}{x^T x} M(x)Y.$$

(Similarly, if v is an eigenvector of Y^TY corresponding to the possible eigenvalue 0, then consideration of Yv shows that 0 is an eigenvalue of $Y - \frac{1}{x^Tx}M(x)Y$.) On the other hand, we have

$$\begin{aligned}(Y - \tfrac{1}{x^Tx}M(x)Y)x &= 0\\ (Y - \tfrac{1}{x^Tx}M(x)Y)Yx &= Y^2x + \lambda_* x = 0.\end{aligned}$$

We see that the eigenvalues of Y other than $\pm{\lambda_*}^{1/2}i$ are also eigenvalues of $Y - \frac{1}{x^Tx}M(x)Y$ with the same multiplicities. It follows that throughout Step 3, the other eigenvalues of Y, and hence Y^TY, will remain constant. (While the preceding discussion seems to assume implicitly that λ_* has multiplicity 1, it is easy to see that the same result applies if there are additional eigenvectors corresponding to λ_*.)

In Step 3, the eigenvalue of Y^TY whose initial value is λ_* is the only one that is evolving nontrivially, and it will in fact converge to 0. To see this, recall that Y converges to $Y_\#$, the projection of Y onto the nullspace of $M(x)$. It follows that

$$\frac{1}{x^Tx}M(x)Y = \frac{1}{x^Tx}(\mu_0 Y_0 + \cdots + \mu_r Y_r)$$

will converge to 0, for if $Y_\# = Y_0 \neq 0$, then $\mu_0 = 0$. But as we have just seen, $\pm{\lambda_*}^{1/2}i$ are the only nonzero eigenvalues of $\frac{1}{x^Tx}M(x)Y$ at time $t = 0$, and thus they must converge to 0 asymptotically.

These considerations also tell us how many times we can expect the stabilization algorithm to iterate. Indeed, since Y^TY can have at most $\lfloor\frac{n}{2}\rfloor$ distinct positive eigenvalues, stabilization will be achieved in at most $\lfloor\frac{n}{2}\rfloor$ iterations.

Specializing further, let us consider the case $so(3)$. Here Y^TY has only one nonzero eigenvalue, which has multiplicity 2. It follows that after one execution of Step 3, Y will converge to 0. Thus the algorithm will stabilize the system with just one iteration of the while loop.

As a numerical example of this, consider the 6th order system

$$\dot{x} = u \tag{6.3.57}$$
$$\dot{Y} = xu^T - ux^T \tag{6.3.58}$$

where $x, u \in R^3$, $Y \in so(3)$ with the following initial conditions:

$$x(0) = \begin{pmatrix} 0.2 \\ 1.1 \\ 1.1 \end{pmatrix}$$

$$Y(0) = \begin{pmatrix} 0 & 0.1 & -0.2 \\ -0.1 & 0 & 3.0 \\ 0.2 & -3.0 & 0 \end{pmatrix}.$$

The spectrum of $Y(0)^T Y(0)$ is $\{9.05, 9.05, 0\}$.

After we apply Step 1 (with $u = -\lambda_* x + Y^T Y x$) over the interval $[0, t_1]$ with $t_1 = 2$ sec, x and Y become respectively

$$x(t_1) = \begin{pmatrix} -0.1076 \\ 1.0966 \\ 1.0726 \end{pmatrix}$$

$$Y(t_1) = \begin{pmatrix} 0 & 0.4359 & 0.1346 \\ -0.4359 & 0 & 2.9738 \\ -0.1346 & -2.9738 & 0 \end{pmatrix}.$$

The spectrum of $Y(t_1)^T Y(t_1)$ remains constant $\{9.05, 9.05, 0\}$ with good accuracy, but as expected, the vector $x(t_1)$ is now the eigenvector corresponding to the eigenvalue $\lambda = 9.05$.

Application of Step 3 (with $u = Yx$) over the interval $[t_1, t_2]$ ($t_2 = 4$ sec) results in the fast decay of $\|Y(t)\|$ to zero as $\|x(t)\|$ remains constant. At the end of this interval, the eigenvalues of $Y(t_2)^T Y(t_2)$ become very small: $\{0.0007, 0.0007, 0\}$.

Finally, Step 4 is executed (with $u = -x$). As expected, x converges to 0 as the value of Y remains unchanged.

Time plots of $\|x(t)\|^2$ and $\|Y(t)\|^2$ are shown in Figure 6.3.1. The decrease in these magnitudes is clearly seen. This provides an illustration of the discontinuous switching control for stabilization developed in §6.3.

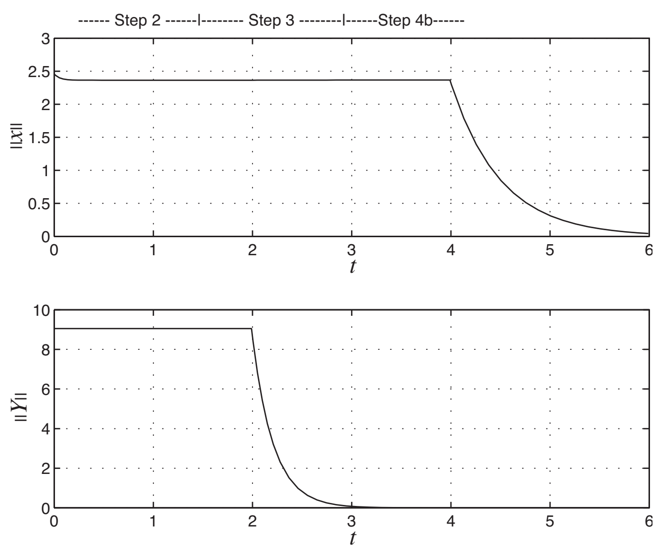

FIGURE 6.3.1. Numerical Example.

6.6 Nonsmooth Stabilization

Control of the Knife Edge Using Steering and Pushing Inputs. Here we describe a further application of the nonsmooth stabilization algorithm of §6.3 of the test (see also Bloch, Reyhanoglu, and McClamroch [1992].) We consider the control of a knife edge moving in point contact on a plane surface. Let x and y denote the coordinates of the point of contact of the knife edge on the plane and let φ denote the heading angle of the knife edge, measured from the x-axis. Then the equations of motion, with all numerical constants set to unity, are given by

$$\ddot{x} = \lambda \sin\varphi + u_1 \cos\varphi \ , \tag{6.6.14}$$
$$\ddot{y} = -\lambda \cos\varphi + u_1 \sin\varphi \ , \tag{6.6.15}$$
$$\ddot{\varphi} = u_2 \ , \tag{6.6.16}$$

where u_1 denotes the control force in the direction defined by the heading angle, u_2 denotes the control torque about the vertical axis through the point of contact; the components of the force of constraint arise from the scalar nonholonomic constraint

$$\dot{x}\sin\varphi - \dot{y}\cos\varphi = 0 \tag{6.6.17}$$

which has nonholonomy degree two at any configuration. It is clear that the constraint manifold is a five-dimensional manifold and is defined by

$$\mathbf{M} = \{(\varphi, x, y, \dot{\varphi}, \dot{x}, \dot{y}) | \dot{x}\sin\varphi - \dot{y}\cos\varphi = 0\}$$

and any configuration is an equilibrium if the controls are zero.

Define the variables $z_1 = x\cos\varphi + y\sin\varphi\,, z_2 = \varphi\,, z_3 = -x\sin\varphi + y\cos\varphi\,, z_4 = \dot{x}\cos\varphi + \dot{y}\sin\varphi - \dot{\varphi}(x\sin\varphi - y\cos\varphi)\,, z_5 = \dot{\varphi}\,,$ so that the reduced differential equations are given by

$$\dot{z}_1 = z_4 \,, \tag{6.6.18}$$
$$\dot{z}_2 = z_5 \,, \tag{6.6.19}$$
$$\dot{z}_3 = -z_1 z_5 \,, \tag{6.6.20}$$
$$\dot{z}_4 = u_1 + u_2 z_3 - z_1 z_5^2 \,, \tag{6.6.21}$$
$$\dot{z}_5 = u_2 \,. \tag{6.6.22}$$

We have:

6.6.4 Proposition. *Let $z^e = (z_1^e, z_2^e, z_3^e, 0, 0)$ denote an equilibrium solution of the reduced differential equations corresponding to $u = 0$. The knife edge dynamics described by equations (6.6.14)-(6.6.15) have the following properties:*

1. *There is a smooth feedback which asymptotically stabilizes the closed loop to any smooth one dimensional equilibrium manifold in $\mathbf{M}$ which satisfies the transversality condition.*

2. *There is no smooth feedback which asymptotically stabilizes* z^e*.*
3. *The system is strongly accessible at* z^e *since the space spanned by the vectors*
$$g_1,\ g_2,\ [g_1, f],\ [g_2, f],\ [g_2, [f, [g_1, f]]]$$
has dimension 5 at z^e*.*
4. *The system is small time locally controllable at* z^e *since the brackets satisfy sufficient conditions for small time local controllability.*

Note that the base variables are (z_1, z_2). Consider a parameterized rectangular closed path γ in the base space with four corner points of the form

$$(0,0),\ (z_1, 0),\ (z_1, z_2),\ (0, z_2)\ ,$$

i.e., $a = (z_1, 0)$ and $b = (0, z_2)$ following the notation introduced in the general development. By evaluating the holonomy integral in closed form for this case, the holonomy equation is

$$z_3^T = z_1 z_2\ .$$

This equation can be explicitly solved (inverted) to determine a closed path $\gamma^* = \gamma(a^*, b^*)$ which achieves the desired holonomy. One solution can be given as follows

$$a^* = \left(\sqrt{|z_3^T|}\,\mathrm{sign}\, z_3^T, 0\right),\ b^* = \left(0, \sqrt{|z_3^T|}\right)\ .$$

Note that the previously described feedback algorithm can be used to asymptotically stabilize the knife edge to the origin.

7

Optimal Control

7.7 Optimal Control on Lie Algebras and Adjoint Orbits

An interesting class of optimal control problems are those which lie naturally on adjoint orbits of compact Lie groups. Analysis of problems of this type may be found in Brockett [1994], Bloch and Crouch [1995, 1996], Bloch, Brockett, and Crouch [1997] and Bloch, Crouch, Marsden, and Ratiu[1998, 2002].

In this section we will set up the variational problem on the adjoint orbits of compact Lie groups and derive the corresponding Hamiltonian equations.

Let G be a compact Lie group (e.g. $SO(n)$) and $\mathfrak{g}$ its Lie algebra. In this case a natural drift free control system on an adjoint orbit of G takes the form

$$\dot{x} = [x, u] \tag{7.7.1}$$

(An arbitrary tangent vector to the adjoint orbit at the point x is of the form $[x, u]$.)

Let the pairing between vectors x in $\mathfrak{g}$ and dual vectors p in $\mathfrak{g}^*$ be written $\langle p, x\rangle = -\kappa(x, p)$ where κ is the Killing form. (We choose the negative sign here to ensure positive definiteness.)

7.7.1 Definition. *Let $T > 0$, $x_0, x_T \in \mathcal{O}$, the orbit in $\mathfrak{g}$ through x_0, be given and fixed. Then we define the optimal control problem*

$$\min_{u \in \mathfrak{g}} \int_0^T \frac{1}{2}||u||^2 - V(x)dt \tag{7.7.2}$$

where $||\cdot|| = \langle \cdot, \cdot \rangle^{1/2}$ is the norm induced on $\mathfrak{g}$ by the negative of the Killing form on $\mathfrak{g}$, V is a smooth function on $\mathfrak{g}$, subject to the constraint on u that there be a curve $x(t) \in \mathfrak{g}$ such that

$$\dot{x} = [x, u] \qquad x(0) = x_0, \qquad x(T) = x_T. \tag{7.7.3}$$

We have

7.7.2 Theorem. *The equations of the maximum principle for the variational problem with functional (7.7.2) subject to the dynamics (7.7.1) are*

$$\begin{aligned} \dot{x} &= [x, [p, x]] \\ \dot{p} &= [p, [p, x]] - V_x \,. \end{aligned} \tag{7.7.4}$$

Proof. The Hamiltonian is given by

$$H(x, p, u) = \langle p, [x, u] \rangle - \frac{1}{2}||u||^2 + V(x)\,. \tag{7.7.5}$$

Hence,

$$\frac{\partial H}{\partial u} = -\langle [x, p], \cdot \rangle - \langle u, \cdot \rangle$$

and thus the optimal control is given by

$$u^* = [p, x]\,. \tag{7.7.6}$$

Substituting this into H, we find that the Hamiltonian evaluated along the optimal trajectory is given by

$$H^*(p, x) = -\frac{1}{2}\langle x, [p, [p, x]] \rangle + V(x)\,. \tag{7.7.7}$$

Computing

$$\dot{x} = \left(\frac{\partial H^*}{\partial p}\right)^T \qquad\qquad \dot{p} = -\left(\frac{\partial H^*}{\partial x}\right)^T$$

gives the result. ■

A particularly interesting special case of this problem is that of Brockett [1994]. His result in this case is as follows:

7.7.3 Corollary. *The equations of the maximum principle for the variational problem (7.7.2) subject to equations (7.7.1) with* $V(x) = -\frac{1}{2}||[x,n]||^2$ *are*

$$\begin{aligned} \dot{x} &= [x, [p, x] \\ \dot{p} &= [p, [p, x]] - [n, [n, x]] \,. \end{aligned} \tag{7.7.8}$$

The proof of the corollary follows immediately, setting

$$V(x) = \frac{1}{2}\langle x, [n, [n, x]]\rangle.$$

Note that with this functional the equations lie naturally on an adjoint orbit. In addition, these equations are interesting in that the optimal flow may be related to the integrable Toda lattice equations (see below and Brockett [1994].)

We next show that these equations may be recast as the Euler-Lagrange equations found by Brockett.

To do this, we introduce the following notation (see e.g. Bloch, Brockett, and Ratiu [1992]): Let x and l lie in $\mathfrak{g}$. Then x may be decomposed as $x = x^l + x_l$ where $x_l \in \mathrm{Ker}(\mathrm{ad}_l)$ and $x^l \in \mathrm{Im}(\mathrm{ad}_l)$ and where $\mathrm{ad}_x(y) = [x, y]$. Further, given any $l \in \mathfrak{g}$ we may decompose $\mathfrak{g}$ orthogonally relative to $-\kappa(,)$ as $\mathfrak{g}^l \oplus \mathfrak{g}_l$ where $\mathfrak{g}^l = \mathrm{Im}(\mathrm{ad}_l)$ and $\mathfrak{g}_l = \mathrm{Ker}(\mathrm{ad}_l)$.

Now any velocity vector is tangent to the orbit of the adjoint action and hence is of the form $\dot{x} = [x, a]$ for some $x \in \mathfrak{g}$. Hence $\dot{x} \in \mathrm{Im}(\mathrm{ad}_x)$. The inverse of operator ad_x which, following Brockett [1994], we will denote by ad_x^{-1}, is well defined on $\mathrm{Im}(\mathrm{ad}_x)$ and hence on $\dot{x}$.

Then we have

7.7.4 Proposition. *The equations (7.7.8) are equivalent to the Euler-Lagrange equations (Brockett [1994])*

$$\ddot{x} = [\dot{x}, \mathrm{ad}_x^{-1}(\dot{x})] + \mathrm{ad}_x^2 \mathrm{ad}_n^2(x). \tag{7.7.9}$$

Proof. Eliminate p from the two equations (7.7.8). The computation is straightforward but somewhat lengthy. ■

The proof of this proposition may of course be obtained also by the Legendre transformation. In terms of the operator ad, the integrand of (7.7.2) with the given $V(x)$ may be written as the Lagrangian (see Brockett [1994])

$$L(x, \dot{x}) = \frac{1}{2}(||[\mathrm{ad}_x^{-1}\dot{x}||^2 + ||[x, n]||^2). \tag{7.7.10}$$

Then, since $\frac{1}{2}||\mathrm{ad}_x^{-1}\dot{x}||^2 = -\frac{1}{2}\langle \mathrm{ad}_x^{-2}\,\dot{x}, \dot{x}\rangle$ we have

$$p = \frac{\partial L}{\partial \dot{x}} = -\mathrm{ad}_x^{-2}\dot{x}.$$

Substituting in (7.7.10) gives the (optimal) Hamiltonian H^*.

We remark that the kinetic energy in (7.7.10) is given by the so-called normal metric (see e.g. Bloch, Brockett, and Ratiu [1992]). This is defined as follows: Let $\mathcal{O}$ be an adjoint orbit of $\mathfrak{g}$ and suppose $\xi = [x,a]$ and $\eta = [x,b]$ are tangent vectors to the orbit at x then $g_n(\xi,\eta) = \langle \xi, \eta \rangle = -\kappa(a^x, b^x)$ where a^x and b^x lie in $\mathrm{Im}(\mathrm{ad}_x)$ as defined above.

We consider now the precise sense in which the equations discussed above are Hamiltonian. The discussion here is brief—more detail may be found in Bloch, Brockett, and Crouch [1997].

We have

7.7.5 Theorem. *Let ω be the standard symplectic structure on $T^*\mathfrak{g}$. Consider the Hamiltonian*

$$H(x,p) = \frac{1}{2}||[p,x]||^2 + V(x) \tag{7.7.11}$$

where V(x) is any smooth function on $\mathfrak{g}$ and p is a momentum variable viewed as lying in $\mathfrak{g}$ by indentifying $\mathfrak{g}$ with its dual. The Hamiltonian equations of motion are

$$\begin{aligned} \dot{x} &= [x,[p,x]] \\ \dot{p} &= [p,[p,x]] - V_x. \end{aligned} \tag{7.7.12}$$

Proof. Let $\xi = (\delta x, \delta p)$ denote an arbitrary tangent vector to $T^*\mathfrak{g}$ and denote the Hamiltonian vector field corresponding to H by $X_H = (\zeta_x, \zeta_p)$. We need to solve for X_H from the equation $dH \cdot \xi = \omega(X_H, \xi)$. Now

$$dH \cdot \xi = \langle [p,x], [\delta p, x] \rangle + \langle [p,x], [p, \delta x] \rangle + \left\langle \frac{\partial V}{\partial x}, \delta x \right\rangle$$

and

$$\omega(X_H, \xi) = \langle \zeta_x, \delta p \rangle - \langle \zeta_p, \delta x \rangle.$$

Equating these expressions gives the result. ∎

We now have

7.7.6 Corollary. *Let $V(x) = \frac{1}{2}\langle [x,n],[x,n] \rangle$ in the Hamiltonian (7.7.11). Then the Hamiltonian equations (7.7.12) yield the optimal Hamiltonian equations (7.7.8).*

Consider now the case $V = 0$. Remarkably, even though these equations are Hamiltonian with respect to the standard symplectic structure on $T^*\mathfrak{g}$, they are in fact Hamiltonian on the cotangent bundle of an adjoint orbit of G. That is, even though the Hamiltonian structure on the orbit is complicated and is not the restriction of the structure on the Lie algebra, the equations themselves do restrict. We have

7.7.7 Theorem. *For $V(x) = 0$ the equations*

$$\begin{aligned} \dot{x} &= [x,[p,x]] \\ \dot{p} &= [p,[p,x]]\,. \end{aligned} \tag{7.7.13}$$

are the Hamiltonian form of the geodesic equations with respect to the normal metric on an adjoint orbit of $\mathfrak{g}$.

Sketch of Proof. From the optimal control calculation above we know that these are the equations of the geodesic flow. It remains to observe that this is the geodesic flow with respect to the normal metric. Now for $V = 0$ the Hamiltonian is just the norm of the velocity in the normal metric. We now need to check that the equations of motion are Hamiltonian with respect to this Hamiltonian and a symplectic structure on the cotangent bundle of the orbit, which may be identified with the tangent bundle. A tangent vector to the tangent bundle to the orbit at the point $(x,[x,\xi])$ is of the form

$$(x,[x,\xi],[x,\eta],[[x,\eta],\xi]+[x,\zeta]) \tag{7.7.14}$$

for $\xi,\eta,\zeta \in \mathfrak{g}$. Then, using the natural symplectic structure as in Thimm [1981] gives the result. The details of this standard but lengthy computation are given in Bloch, Brockett, and Crouch [1997]. ■

We may also endow any orbit with the right invariant metric

$$g_{nl}([x,a],[x,b]) = -\kappa(a^x, Jb^x) \tag{7.7.15}$$

where J is a positive self-adjoint operator on the algebra. Then we have

7.7.8 Corollary. *The geodesic equations on an adjoint orbit endowed with the right invariant metric* (7.7.15) *are*

$$\begin{aligned} \dot{p} &= [p,J^{-1}[p,x]] \\ \dot{x} &= [x,J^{-1}[p,x]]\,. \end{aligned} \tag{7.7.16}$$

We shall consider this right invariant case from the optimal control point of view in the next paragraph.

We note also that, as expected, in the bi-invariant case with $V = 0$ it is possible to explicitly compute the solutions of (7.7.13) despite their strong coupling. This follows from

7.7.9 Lemma. $[p,x]$ *is conserved along the flow of* (7.7.13).

This is proved by a simple computation. However proving complete integrability in the Hamiltonian sense, i.e., finding a complete set of commuting integrals, is by no means easy. This is the content of Thimm [1981] and Bloch, Brockett, and Crouch [1997].

Optimal Control on Symmetric Spaces. The equations discussed above are not only well defined on adjoint orbits but also on general symmetric spaces where the tangent vectors to the space are given in the form a suitable bracket—this includes the complex and real Grassmannians of q-planes in $n+1$-space $G_{q,n+1}(\mathbb{C})$ or $G_{q,n+1}(\mathbb{R})$ and in particular the spheres.

This may be seen as follows: The complex Grassmannian is given by

$$U(n+1)/U(q)\times U(p), \qquad q+p=n+1,\, q\le p \tag{7.7.17}$$

and the real Grassmannian by

$$\mathrm{SO}(n+1)\,\mathrm{SO}(q)\times \mathrm{SO}(p), \qquad q+p=n+1,\, q\le p \tag{7.7.18}$$

where $U(n)$ is the unitary group and $\mathrm{SO}(n)$ the special orthogonal group. In either case let $\mathfrak{g}=\mathfrak{k}\oplus\mathfrak{m}$ be the Lie algebra decomposition corresponding to G/K. We may thus represent a point in the complex (real) Grassmannian by a matrix

$$\hat{Q}=\begin{bmatrix} 0 & Q \\ -Q^* & 0 \end{bmatrix} \tag{7.7.19}$$

in $\mathfrak{m}$ where Q is a complex (real) $p\times q$ matrix of full rank and Q^* is its Hermitian conjugate (transpose). A point in $\mathfrak{k}$ may be represented by the matrix

$$\hat{K}=\begin{bmatrix} K_1 & O \\ O & K_2 \end{bmatrix} \tag{7.7.20}$$

where $K_1\in u(p)(so(p))$ and $K_2\in u(q)(so(q))$. Define $\hat{P}$ to be a similarly partitioned matrix. Then we have

7.7.10 Proposition. *Tangent vectors to the Grassmannian are represented by matrices of the form*

$$[\hat{Q},\hat{K}]\,.$$

Proof. A curve in the Grassmannian through the point $\hat{Q}$ may be given by

$$e^{-\hat{K}t}\hat{Q}e^{\hat{K}t}.$$

Note that the given curve simply provides an orthogonal (or unitary) transformation of the rows and columns of Q. Differentiating at $t=0$ gives the result. ■

Since tangent vectors are given by brackets, just as in the case of orbits, a normal metric may be defined as follows: Given two tangent vectors $\xi_1=[\hat{Q},\hat{K}_1]$ and $\xi_2=[\hat{Q},\hat{K}_2]$ define $<\xi_1,\xi_2>_N$ to be $-K(\hat{K}_1^{\hat{Q}},\hat{K}_2^{\hat{Q}})$ where $\hat{K}_1^{\hat{Q}}$ denotes the projection of $\hat{K}_1$ onto the image of $\mathrm{ad}_{\hat{Q}}$ in the decomposition $ls=\mathrm{Ker}(\mathrm{ad}_{\hat{Q}})\oplus\mathrm{Im}(\mathrm{ad}_{\hat{Q}})$. As in the case of orbits we thus obtain

7.7.11 Proposition. *The geodesic equations on the real or complex Grassmannian are given by*

$$\begin{aligned}\dot{\hat{Q}} &= [\hat{Q}, [\hat{P}, \hat{Q}]] \\ \dot{\hat{P}} &= [\hat{P}, [\hat{P}, \hat{Q}]]\,.\end{aligned} \tag{7.7.21}$$

where $\hat{Q}$ is given by 7.7.19 and similarly for $\hat{P}$.

For a symmetric space we have $[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}$, $[\mathfrak{k}, \mathfrak{m}] \subset \mathfrak{m}$. Thus, $[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}$ and since $\hat{Q}, \hat{P} \in \mathfrak{m}$ the equations are naturally well defined.

In fact the formalism developed here can be combined with the work of Thimm [1981] to give an explicit proof of complete integrability of the geodesic flow on symmetric spaces such as the real and complex Grassmannians. In particular it is possible to derive explicitly a complete set of commuting flows and to prove their involutivity. This is the subject of Bloch, Brockett, and Crouch [1997].

This should be contrasted with the observation of Brockett regarding the optimal control equations for (7.7.2) with $V(x) = \frac{1}{2}||[x, n]||^2$. He observed that the optimal flow under suitable conditions was given by $\dot{x} = [x, [x, n]]$ and in the particular case of x being tridiagonal one obtains the integrable Toda lattice equations in Flaschka's form but written as a double bracket equation (see Bloch [1990]). Integrability of optimal control problems is a rich subject (see, for example, Faybusovich [1988]).

As in the variational problem on adjoint orbits the double, double bracket equations on Grassmannians are solutions to a natural optimal problem.

7.7.12 Definition. *Let $\mathfrak{u}(n)$ denote the Lie algebra of the unitary group* $\mathrm{U}(n)$. *Let Q be a $p \times q$ complex matrix and let $U \in \mathfrak{u}(p)$ and $V \in \mathfrak{u}(q)$. Let J_U and J_V be constant symmetric positive definite operators on the space of complex $p \times p$ and $q \times q$ matrices respectively and let $\langle \cdot, \cdot \rangle$ denote the trace inner product $\langle A, B \rangle = \frac{1}{2}\,\mathrm{trace}(A^\dagger B)$, where $A^\dagger$ is the adjoint; that is, the transpose conjugate.*

Let $T > 0$, Q_0, Q_T be given and fixed. Define the optimal control problem over $\mathfrak{u}(p) \times \mathfrak{u}(q)$

$$\min_{U,V} \frac{1}{4} \int \{\langle U, J_U U \rangle + \langle V, J_V V \rangle\} dt \tag{7.7.22}$$

subject to the constraint that there exists a curve $Q(t)$ such that

$$\dot{Q} = UQ - QV, \qquad Q(0) = Q_0, \qquad Q(T) = Q_T. \tag{7.7.23}$$

7.7.13 Theorem. *The optimal control problem (7.7.12) has optimal controls given by*

$$\begin{aligned}U &= J_U^{-1}(PQ^\dagger - QP^\dagger) \\ V &= J_V^{-1}(P^\dagger Q - Q^\dagger P)\,.\end{aligned} \tag{7.7.24}$$

and the optimal evolution of the states Q and costates P is given by

$$\begin{aligned}\dot{Q} &= J_U^{-1}(PQ^\dagger - QP^\dagger)Q - QJ_V^{-1}(P^\dagger Q - Q^\dagger P)\\ \dot{P} &= J_U^{-1}(PQ^\dagger - QP^\dagger)P - PJ_V^{-1}(P^\dagger Q - Q^\dagger P).\end{aligned} \tag{7.7.25}$$

Note: J_U and J_V are in general different operators acting on different spaces. In certain case (see the rigid body below) the spaces and the operators may be taken to be the same.

Proof. Form the Hamiltonian

$$H(Q,P,U,V) = \langle P, UQ - QV\rangle - \frac{1}{2}\langle U, J_U U\rangle - \frac{1}{2}\langle V, J_V V\rangle\,. \tag{7.7.26}$$

To find the optimal control we differentiate the Hamiltonian with respect to U and V in the directions Y and Z respectively and set equal to zero. This yields

$$\langle P, YQ\rangle - \frac{1}{4}\langle Y, J_U U\rangle - \frac{1}{4}\langle U, J_U Y\rangle = 0$$

and

$$\langle P, -QZ\rangle - \frac{1}{4}\langle Z, J_V V\rangle - \frac{1}{4}\langle V, J_V Z\rangle = 0\,.$$

We have $\overline{\langle Y, J_U U\rangle} = \langle U, J_U Y\rangle$ and $\overline{\langle U, J_U Y\rangle} = \langle Y, J_U U\rangle$ and thus along the optimal trajectory

$$\overline{\langle P, YQ\rangle} = \langle P, YQ\rangle\,. \tag{7.7.27}$$

Now

$$\langle P, YQ\rangle = \frac{1}{2}\overline{\langle Y, PQ^\dagger\rangle} - \frac{1}{2}\langle Y, QP^\dagger\rangle$$

and

$$\overline{\langle P, YQ\rangle} = \frac{1}{2}\langle Y, PQ^\dagger *\rangle - \frac{1}{2}\overline{\langle Y, QP^\dagger\rangle}\,.$$

Thus, using the fact that $\langle P, YQ\rangle$ is real, along the optimal trajectory we have

$$\begin{aligned}\langle P, YQ\rangle &= \frac{1}{2}\langle P, YQ\rangle + \frac{1}{2}\overline{\langle P, YQ\rangle}\\ &= \frac{1}{4}\overline{\langle Y, PQ^\dagger - QP^\dagger\rangle} + \frac{1}{4}\langle Y, PQ^\dagger - QP^\dagger\rangle\\ &= \frac{1}{4}\langle Y, J_U U\rangle + \frac{1}{4}\overline{\langle Y, J_U U\rangle}\,.\end{aligned} \tag{7.7.28}$$

Hence

$$J_U U = PQ^\dagger - QP^\dagger.$$

Similarly for V.

Now the equations for Q and P are given by

$$\begin{aligned} \langle Z, \dot{Q}\rangle &= \nabla_P H(Z) = \langle Z, UQ - QV\rangle \\ \langle \dot{P}, Z\rangle &= -\nabla_Q H(Z) = -\langle P, UZ - ZV\rangle = \langle UP - PV, Z\rangle \,, \end{aligned} \tag{7.7.29}$$

and hence the result. ■

We remark that this result does not preclude the existence of conjugate points.

We have the immediate corollary:

7.7.14 Corollary. *For J_U and J_V equal to the identity, the optimal control equations for the problem (7.7.22) subject to (7.7.23) are*

$$\begin{aligned} \dot{Q} &= PQ^\dagger Q + QQ^\dagger P - 2QP^\dagger Q \\ \dot{P} &= 2PQ^\dagger P - QP^\dagger P - PP^\dagger Q \,. \end{aligned} \tag{7.7.30}$$

Further, in the general case we have

7.7.15 Corollary. *The equations (7.7.25) are given by the double double bracket equations*

$$\begin{aligned} \dot{\hat{Q}} &= [\hat{Q}, \hat{J}^{-1}[\hat{P}, \hat{Q}]] \\ \dot{\hat{P}} &= [\hat{P}, \hat{J}^{-1}[\hat{P}, \hat{Q}]] \,. \end{aligned} \tag{7.7.31}$$

where $\hat{J}$ is the operator $\operatorname{diag}(J_U, J_V)$,

$$\hat{Q} = \begin{bmatrix} 0 & Q \\ -Q^\dagger & 0 \end{bmatrix} \in \mathfrak{u}(p+q), \tag{7.7.32}$$

Q is a complex $p \times q$ matrix of full rank, $Q^\dagger$ is its adjoint, and similarly for P.

The proof is a computation.

Note that the equations (7.7.31) or (7.7.25) give the geodesic equations on the complex Grassmannian (or real Grassmannian in the real case) with respect to the right invariant "normal" metric.

As an example in the current setting, we next write explicitly the geodesic flow on the sphere S^n. Recall (see e.g. Moser [1980]) that the geodesic motion on S^n may be written as follows: Let $\mathbf{q} = [q_1, \cdots, q_{n+1}]^T \in \mathbb{R}^{n+1}$ with Euclidean norm $||\mathbf{q}|| = 1$, represent an element of S^n. Then the geodesic flow can be found by setting $\ddot{\mathbf{q}} = \lambda \mathbf{q}$ where λ is chosen so that $||\mathbf{q}||$ is compatible with the flow. This follows from the constrained optimization of the free particle Lagrangian. This implies $\langle \mathbf{q}, \dot{\mathbf{q}}\rangle = 0$ and $\langle \mathbf{q}, \ddot{\mathbf{q}}\rangle + ||\dot{\mathbf{q}}||^2 = 0$. Thus $\lambda = -||\dot{\mathbf{q}}||^2$ and the geodesic flow is given by

$$\ddot{\mathbf{q}} = -||\dot{\mathbf{q}}||^2 \mathbf{q}. \tag{7.7.33}$$

Letting $\mathbf{p} = [p_1, \cdots, p_{n+1}]^T \in \mathbb{R}^{n+1}$, this may be viewed as a Hamiltonian system restricted to $||\mathbf{q}|| = 1, \langle \mathbf{q}, \mathbf{p} \rangle = 0$. With Hamiltonian $H = \frac{1}{2}||\mathbf{q}||^2||\mathbf{p}||^2$, we get the flow

$$\dot{\mathbf{q}} = \left(\frac{\partial H}{\partial \mathbf{p}}\right)^T = \mathbf{p} \qquad \dot{\mathbf{p}} = -\left(\frac{\partial H}{\partial \mathbf{q}}\right)^T = -||\mathbf{p}||^2 \mathbf{q}\,. \tag{7.7.34}$$

In our current setting we have

7.7.16 Proposition. *Let*

$$\hat{P} = \begin{bmatrix} 0 & \cdots & 0 & p_1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & p_{n+1} \\ -p_1 & \cdots & -p_{n+1} & 0 \end{bmatrix} \qquad \hat{Q} = \begin{bmatrix} 0 & \cdots & 0 & q_1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & q_{n+1} \\ -q_1 & \cdots & -q_{n+1} & 0 \end{bmatrix} \tag{7.7.35}$$

where we normalize $||\mathbf{q}|| = 1$ *and* $\langle \mathbf{q}, \mathbf{p} \rangle = 0$. *Then the flow* (7.7.13) *yields the geodesic flow* (7.7.33).

The proof is a straightforward computation.

Exercises

◇ **7.7-1.** Formulate the geodesic flow on a sphere directly as an optimal control problem

This flow is completely integrable and again provides an example of an integrable optimal control problem. Details of the proof of integrability can be found in Thimm [1981] or Bloch, Brockett, and Crouch [1997].

The rigid body as an optimal control problem.[1]

We now discuss the SO(n) rigid body equations in the optimal control setting.

For convenience we shall use the following pairing (multiple of the Killing form) on $\mathfrak{so}(n)$, the Lie algebra of $n \times n$ real skew matrices regarded as the Lie algebra of the n-dimensional proper rotation group SO(n):

$$\langle \xi, \eta \rangle = -\frac{1}{2}\,\mathrm{trace}(\xi\eta).$$

The factor of 1/2 in this formula is to make this inner product agree with the usual inner product on $\mathbb{R}^3$ when it is identified with $\mathfrak{so}(3)$ in the following standard way: we associate the 3×3 skew matrix $\hat{u}$ to the vector u by $\hat{u} \cdot v = u \times v$, where $u \times v$ is the usual cross product in $\mathbb{R}^3$.

[1]Based on Bloch, Crouch, Marsden, and Ratiu [2002]

We use this inner product to identify the dual of the Lie algebra, namely $\mathfrak{so}(n)^*$, with the Lie algebra $\mathfrak{so}(n)$.

We recall from Manakov [1976] and Ratiu [1980] that the left invariant generalized rigid body equations on $\mathrm{SO}(n)$ may be written as

$$\begin{aligned} \dot{Q} &= Q\Omega \\ \dot{M} &= [M, \Omega], \end{aligned} \tag{7.7.36}$$

where $Q \in \mathrm{SO}(n)$ denotes the configuration space variable (the attitude of the body), $\Omega = Q^{-1}\dot{Q} \in \mathfrak{so}(n)$ is the body angular velocity, and $M := J(\Omega) \in \mathfrak{so}(n)$ is the body angular momentum. Here $J : \mathfrak{so}(n) \to \mathfrak{so}(n)$ is the symmetric (with respect to the above inner product) positive definite operator defined by

$$J(\Omega) = \Lambda\Omega + \Omega\Lambda,$$

where Λ is a diagonal matrix satisfying $\Lambda_i + \Lambda_j > 0$ for all $i \neq j$. For $n = 3$ the elements of Λ_i are related to the standard diagonal moment of inertia tensor I by $I_1 = \Lambda_2 + \Lambda_3$, $I_2 = \Lambda_3 + \Lambda_1$, $I_3 = \Lambda_1 + \Lambda_2$.

The equations $\dot{M} = [M, \Omega]$ are readily checked to be the Euler-Poincaré equations on $\mathfrak{so}(n)$ for the Lagrangian

$$l(\Omega) = \frac{1}{2}\langle \Omega, J(\Omega) \rangle .$$

It follows from general Euler-Poincaré theory (see, for example, Marsden and Ratiu [1999]) that the equations (7.7.36) are the geodesic equations on $T\,\mathrm{SO}(n)$, left trivialized as $\mathrm{SO}(n) \times \mathfrak{s}o(n)$, relative to the left invariant metric whose expression at the identity is

$$\langle\langle \Omega_1, \Omega_2 \rangle\rangle = \langle \Omega_1, J(\Omega_2) \rangle . \tag{7.7.37}$$

According to Mishchenko and Fomenko [1978], there is a similar formalism for any semisimple Lie group and that in that context, one has integrability on the generic coadjoint orbits.

The Symmetric Rigid Body System. By definition, ***the left invariant symmetric rigid body system*** is given by the first order equations

$$\begin{aligned} \dot{Q} &= Q\Omega \\ \dot{P} &= P\Omega \end{aligned} \tag{7.7.38}$$

where Ω is regarded as a function of Q and P via the equations

$$\Omega := J^{-1}(M) \in \mathfrak{so}(n) \quad \text{and} \quad M := Q^T P - P^T Q.$$

It is easy to check that this system of equations on the space $\mathrm{SO}(n) \times \mathrm{SO}(n)$ is invariant under the left diagonal action of $\mathrm{SO}(n)$.

7.7.17 Proposition. *If (Q, P) is a solution of (7.7.38), then (Q, M) where $M = J(\Omega)$ and $\Omega = Q^{-1}\dot{Q}$ satisfies the rigid body equations (7.7.36).*

Proof. Differentiating $M = Q^T P - P^T Q$ and using the equations (7.7.38) gives the second of the equations (7.7.36). ■

It is because of this proposition that the equations (7.7.38) are called the left invariant symmetric rigid body equations on $\mathrm{SO}(n) \times \mathrm{SO}(n)$.

Recall that the ***spatial angular momentum*** for the standard left invariant rigid body equations (7.7.36) is *defined* to be the value of momentum map for the cotangent lifted left action of $\mathrm{SO}(n)$ on $T^* \mathrm{SO}(n)$ (see, for example, Marsden and Ratiu [1999] for these basic notions).

7.7.18 Proposition. *For a solution of the left invariant rigid body equations (7.7.36) obtained by means of Proposition 7.7.17, the spatial angular momentum is given by $m = PQ^T - QP^T$ and hence m is conserved along the rigid body flow.*

Proof. If we start with a solution $(Q(t), P(t))$ of the symmetric rigid body system, and map this solution to $(Q(t), M(t))$ where $M(t) = Q^T P - P^T Q$, then as we have seen, M satisfies the rigid body system, and so M is the body angular momentum, that is, it is the value of the momentum map for the *right* action. By general Euler-Poincaré and Lie-Poisson theory, m, which is the value of the momentum map for the left action, is obtained from M using the coadjoint action of $\mathrm{SO}(n)$ on $\mathfrak{so}(n)^* \cong \mathfrak{so}(n)$, namely

$$m = QMQ^T = Q(Q^T P - P^T Q)Q^T = PQ^T - QP^T.$$

It follows by Noether's theorem that $\dot{m} = 0$; one can also verify this directly by differentiating m along (7.7.38). ■

Local Equivalence of the Rigid Body and the Symmetric Rigid Body Equations. Above we saw that solutions of the symmetric rigid body system can be mapped to solutions of the rigid body system. Now we consider the converse question. Thus, suppose we have a solution (Q, M) of the standard left invariant rigid body equations. We seek to solve for P in the expression

$$M = Q^T P - P^T Q. \tag{7.7.39}$$

For the following discussion, it will be convenient to make use of the operator norm on matrices. We recall, for notational purposes, that this norm is given by

$$\|A\|_{\mathrm{op}} = \sup_{\|x\|=1} \|Ax\|$$

where the norms on the right hand side are the usual Euclidean space norms.

Since elements of $\mathrm{SO}(n)$ have operator norms bounded by 1 and since the operator norm satisfies $\|AB\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}} \|B\|_{\mathrm{op}}$, we see that *if M satisfies $M = Q^T P - P^T Q$, then $\|M\|_{\mathrm{op}} \leq 2$.* Therefore, $\|M\|_{\mathrm{op}} \leq 2$ *is a necessary condition for solvability of* (7.7.39) *for P.*

7.7.19 Definition. *Let C denote the set of (Q, P) that map to M's with operator norm equal to 2 and let S denote the set of (Q, P) that map to M's with operator norm strictly less than 2. Also denote by S_M the set of points $(Q, M) \in T^* \mathrm{SO}(n)$ with $\|M\|_{\mathrm{op}} \leq 2$.*

Note that C contains pairs (Q, P) with the property that $Q^T P$ is both skew and orthogonal.

Recall that $\sinh : \mathfrak{so}(n) \to \mathfrak{so}(n)$ is defined by

$$\sinh \xi = \frac{e^{\xi} - e^{-\xi}}{2}.$$

One sees that indeed sinh takes values in $\mathfrak{so}(n)$ by using, for example, its series expansion:

$$\sinh \xi = \xi + \frac{1}{3!}\xi^3 + \frac{1}{5!}\xi^5 + \dots .$$

Recall from calculus that the inverse function $\sinh^{-1}(u)$ has a convergent power series expansion for $|u| < 1$ that is given by integrating the power series expansion of the function $1/\sqrt{1+u^2}$ term by term. This power series expansion shows that the map $\sinh : \mathfrak{so}(n) \to \mathfrak{so}(n)$ has an inverse on the set $U = \{u \in \mathfrak{so}(n) \mid \|u\|_{\mathrm{op}} < 1\}$. We shall denote this inverse, naturally, by $\sinh^{-1}$, so

$$\sinh^{-1} : U \to \mathfrak{so}(n).$$

Example of $\mathrm{SO}(3)$**.** As an example, let us consider $\mathfrak{so}(3)$ which we parameterize as follows: we write an element of $\mathfrak{so}(3)$ as $\mu\hat{c}$ where $\hat{c}$ is an element of $\mathfrak{so}(3)$ of unit operator norm (so c, the corresponding 3-vector has vector norm one) and μ is a positive scalar. One checks that the operator norm of $\hat{c}$ is equal to the Euclidean norm of c. Hence, the set U consists of the set of elements $\mu\hat{c}$ where c is a unit vector and μ is a real number with $0 \leq \mu < 1$. From Rodrigues' formula (see e.g. Marsden and Ratiu [1999]) one finds that

$$e^{\mu\hat{c}} = I + \sin(\mu)\hat{c} + \left(I - cc^T\right)(\cos\mu - 1). \tag{7.7.40}$$

Thus, one sees that

$$\sinh(\mu\hat{c}) = \sin(\mu)\hat{c}\,.$$

Notice that from this formula, sinh is not globally one to one. However, it has an inverse defined on the set U explicitly given by

$$\sinh^{-1}(\mu\hat{c}) = \sin^{-1}(\mu)\hat{c}.$$

7.7.20 Proposition. *For $\|M\|_{\mathrm{op}} < 2$, the equation (7.7.39) has the solution*

$$P = Q\left(e^{\sinh^{-1} M/2}\right). \tag{7.7.41}$$

Proof. Notice that

$$M = e^{\sinh^{-1} M/2} - e^{-\sinh^{-1} M/2} . \quad \blacksquare$$

Similarly, in the right invariant case, we obtain the formula

$$P_r = \left(e^{\sinh^{-1} M_r/2} \right) Q_r \, . \tag{7.7.42}$$

Example of SO(3)**.** We now show that for $SO(3)$ the set C is not empty, even though there are no points Q, P such that $Q^T P$ is both skew and orthogonal (because in $SO(3)$ there are no skew orthogonal matrices, as all three by three skew matrices are singular). Let $Q^T P = e^{\mu \hat{c}}$ where $\mu = \pi/2$. Then by equation (7.7.40) $Q^T P = I + \hat{c}$ and hence is not skew. Now for x such that $c^T x = 0$ we have

$$\|(Q^T P - P^T Q)x\| = 2\|\hat{c}x\| = 2\|x\|$$

and thus $\|(Q^T P - P^T Q)\|_{\mathrm{op}} = 2$.

In fact, reversing the argument above shows that for $SO(3)$ *the set* C *consists entirely of elements of form* $Q^T P = I + \hat{c}$ *for some* c*.*

7.7.21 Proposition. *The sets* C *and* S *are invariant under the double rigid body equations.*

Proof. Notice that the operator norm is invariant under conjugation; that is, for $Q \in \mathrm{SO}(n)$ and $M \in \mathfrak{so}(n)$, we have

$$\|QMQ^{-1}\|_{\mathrm{op}} = \|M\|_{\mathrm{op}}.$$

This is readily checked from the definition of the operator norm. Recall that under the identification of the dual $\mathfrak{so}(n)^*$ with the space $\mathfrak{so}(n)$, the coadjoint action agrees with conjugation. Thus, the map $f : \mathfrak{so}(3) \to \mathbb{R}$; $M \mapsto \|M\|_{\mathrm{op}}$ is a Casimir function and so is invariant under the dynamics. In particular, its level sets are invariant and so the sets S and C are invariant. $\blacksquare$

One can see that the operator norm is invariant under the dynamics by a direct argument as well. This is done by writing the operator norm as $\|M\|_{\mathrm{op}} = \sqrt{\lambda}$, where λ is the maximum eigenvalue of $M^T M$ (by the Rayleigh-Ritz quotient). Then one differentiates the equation $M^T M v = \lambda v$ along the flow of the rigid body equations, subject to the constraint $\|v\|^2 = 1$ to see that $\dot{\lambda} = 0$.

Example of SO(3)**.** For the rotation group, the trace norm (up to a factor of 2) and the operator norm both agree with the standard Euclidean norm under the identification $v \in \mathbb{R}^3 \mapsto \hat{v} \in \mathfrak{so}(3)$. The standard norm is indeed a Casimir function for the rotation group and is invariant under the rigid body equations by conservation of angular momentum.

We can obtain these equations from the optimal control viewpoint as follows:

7.7.22 Definition. *Let $T > 0$, $Q_0, Q_T \in \mathrm{SO}(n)$ be given and fixed. Let the rigid body optimal control problem be given by*

$$\min_{U \in \mathfrak{so}(n)} \frac{1}{2} \int_0^T \langle U, J(U) \rangle dt \tag{7.7.43}$$

subject to the constraint on U that there be a curve $Q(t) \in \mathrm{SO}(n)$ such that

$$\dot{Q} = QU\,, \qquad Q(0) = Q_0, \qquad Q(T) = Q_T. \tag{7.7.44}$$

7.7.23 Proposition. *The rigid body optimal control problem (7.7.22) has optimal evolution equations (7.7.38) where P is the costate vector given by the maximum principle.*

The optimal controls in this case are given by

$$U = J^{-1}(Q^T P - P^T Q). \tag{7.7.45}$$

Remark The proof (see Bloch and Crouch [1996]) simply involves writing the Hamiltonian of the maximum principle as

$$H = \langle P, QU \rangle - \frac{1}{2} \langle U, J(U) \rangle \tag{7.7.46}$$

and maximizing with respect to U in the standard fashion.

Exercises

◇ **7.7-2.** Show that if instead of the Hamiltonian (7.7.46) one formulates the Hamiltonian in Proposition 7.7.23 as

$$H = \langle QP, QU \rangle - \frac{1}{2} \langle U, J(U) \rangle \tag{7.7.47}$$

where QP is now a left invariant vector field and P is in $so(n)$ one obtains the rigid body equations in the standard rather than symmetric form

Now we show that the rigid body equations may be given as singular case of the double double bracket equations discussed earlier for the general optimal control problem. To see this, let

$$\hat{Q} = \begin{bmatrix} 0 & Q \\ -Q^T & 0 \end{bmatrix} \tag{7.7.48}$$

as before and similarly for $\hat{P}$. Note that these matrices now lie in $\mathfrak{so}(2n)$ and each block lies in $\mathrm{SO}(n)$.

7.7.24 Corollary. *The generalized rigid body equations on* $\mathrm{SO}(n)$ *are given by the double double bracket equations* (7.7.31) *in the case* Q *and* P *lie in* $\mathrm{SO}(n)$, $J_U = J$, *and the operator* $J_V^{-1} = 0$.

Note that the reduced generalized rigid body equations (the dynamics) are completely integrable (see e.g. Ratiu [1980]).

There is also a fascinating symmetric discrete rigid body system which can also be obtained via the (discrete) maximum principle and which is connected with the discrete integrable Moser-Veselov equations for the rigid body Moser and Veselov [1991]. We just mention this briefly here and refer the reader to Bloch, Crouch, Marsden, and Ratiu [1998, 2002] for details.

The Discrete Symmetric Rigid Body. We now define the symmetric discrete rigid body equations as follows:

$$\begin{aligned} Q_{k+1} &= Q_k U_k \\ P_{k+1} &= P_k U_k \, , \end{aligned} \tag{7.7.49}$$

where U_k is defined by

$$U_k \Lambda - \Lambda U_k^T = Q_k^T P_k - P_k^T Q_k = J^D(U_k) \, . \tag{7.7.50}$$

We can then obtain the discrete symmetric rigid body equations as follows:

7.7.25 Definition. *Let* Λ *be a positive definite diagonal matrix. Let* $\overline{Q}_0, \overline{Q}_T \in \mathrm{SO}(n)$ *be given and fixed. Define the optimal control problem*

$$\min_{U_k} \sum_k \operatorname{trace}(\Lambda U_k) \tag{7.7.51}$$

subject to dynamics and initial and final data

$$Q_{k+1} = Q_k U_k, \qquad Q_0 = \overline{Q}_0, \qquad Q(T) = \overline{Q}_T \tag{7.7.52}$$

for $Q_k, U_k \in \mathrm{SO}(n)$.

7.7.26 Theorem. *The optimal control problem* 7.7.25 *yields the optimal evolution equations*

$$\begin{aligned} Q_{k+1} &= Q_k U_k \\ P_{k+1} &= P_k U_k \, , \end{aligned} \tag{7.7.53}$$

where P_k *is the discrete covector and* U_k *is defined by*

$$U_k \Lambda - \Lambda U_k^T = Q_k^T P_k - P_k^T Q_k \, . \tag{7.7.54}$$

Optimal Control and the Full Toda Flow.

In this section we introduce an optimal control problem which yields the full Toda flow as described originally in Faybusovich [1988]. We follow here the treatment in Bloch and Crouch [1997]. This problem may be viewed as a control problem on an adjoint orbit of lower triangular matrices.

We begin by introducing some specialized notation. Let $\mathcal{G} = \mathfrak{gl}(n)$ denote the Lie algebra of $n \times n$ matrices (with corresponding Lie group $GL(n)$ of invertible $n \times n$ matrices).

7.7.27 Definition. *For $A \in \mathfrak{gl}(n)$, let $A = A_+ + A_0 + A_-$ where A_+ is the strictly upper part of A, A_0 the diagonal part, and A_- is the strictly lower part. Let*

$$\begin{aligned}
\pi_l(A) &= A_- + A_+^T + A_0 \ \textit{(lower)} \\
\pi_k(A) &= A_+ - A_+^T \ \textit{(skew)} \\
\pi_{k^\perp}(A) &= A_- + A_0 + A_-^T \ \textit{(symmetric)} \\
\pi_{l^\perp}(A) &= A_+ - A_-^T \ \textit{(strictly upper).}
\end{aligned}$$

The reason for the above notation is as follows (see Symes [1980]):

Endow $\mathcal{G}$ with the inner product $\langle A, B\rangle_T = \mathrm{Tr} A^T B$. We observe

$$\mathcal{G} = \mathcal{L} \oplus \mathcal{K}$$

where $\mathcal{L}$ is the subalgebra of lower triangular matrices and $\mathcal{K}$ the subalgebra of skew-symmetric matrices. $\mathcal{L}$ is the Lie algebra of L, the lower triangular group and $\mathcal{K}$ is the Lie algebra of K, the group of orthogonal matrices.

Denote by $\perp$ the perpendicular subspace under the scalar product. Then

$$\mathcal{L}^\perp = \{x : \mathrm{Tr}(x^T y) = 0, \forall y \in \mathcal{L}\}$$

is the algebra of strictly upper triangular matrices and

$$\mathcal{K}^\perp = \{x : \mathrm{Tr}(x^T z) = 0, \forall z \in \mathcal{K}\}$$

is the space of all symmetric matrices. Hence we can write

$$\mathcal{G} \simeq \mathcal{G}^* = \mathcal{L}^\perp \oplus \mathcal{K}^\perp,$$

and we can make the identification

$$\begin{aligned}
\mathcal{L}^* &\simeq \mathcal{K}^\perp \\
\mathcal{K}^* &\simeq \mathcal{L}^\perp.
\end{aligned}$$

We now give some preliminary results that will be useful in the main theorem.

7.7.28 Lemma. *For any matrix $S \in \mathcal{G}$ and L in$\mathcal{L}$*

$$\begin{aligned}\pi_{k^\perp}(L^T S) &= \pi_{k^\perp}(L^T \pi_{k^\perp} S)\\ \pi_{k^\perp}(SL^T) &= \pi_{k^\perp}(\pi_{k^\perp}(S) L^T).\end{aligned}$$

Let $F : \mathcal{G} \to \mathcal{G}$ be analytic and consider

$$f(PQ^T) = \mathrm{Tr}F(PQ^T).$$

Let

$$\mathcal{F}_A f(A)(R) = \lim_{h\to 0} \frac{f(A+hR) - f(A)}{h} = \langle R, \nabla f(A)\rangle$$

for $A, R \in gl(n)$. Hence

$$\begin{aligned}\mathcal{F}_Q f(PQ^T)(R) &= \langle PR^T, \nabla f(PQ^T)\rangle\\ &= \langle RP^T, \nabla f(PQ^T)^T\rangle\\ &= \langle R, \nabla f(PQ^T)^T P\rangle\\ \mathcal{F}_P f(PQ^T)(R) &= \langle RQ^T, \nabla f(PQ^T)\rangle\\ &= \langle R, \nabla f(PQ^T) Q\rangle.\end{aligned}$$

Hence, we have

7.7.29 Lemma. *Let $f(PQ^T) = TrF(PQ^T)$. Then the Hamiltonian flows with respect to the canonical structure on $T^*\mathcal{G}$ are given by*

$$\begin{aligned}\dot{Q} &= \nabla f(PQ^T) Q\\ \dot{P} &= -\nabla f(PQ^T)^T P,\end{aligned}$$

respectively.

The main result in Bloch and Crouch [1997] can now be stated in the following way. Consider the optimal control problem

$$\begin{aligned}&\min_U \int_0^T \tfrac{1}{2}\langle U, U\rangle_T dt\\ &\text{subject to: } \dot{X} = \pi_l(U)X,\ X(0) = X_0,\ X(T) = X_T,\\ &\text{where } X \in L,\ U \in \mathcal{G}.\end{aligned} \tag{7.7.55}$$

7.7.30 Theorem. *For the optimal control problem (7.7.55), the optimal controls are given by*

$$U = \pi_{k^\perp}(RX^T)$$

where R is the costate vector and the corresponding extremal flow is given by

$$\begin{aligned}\dot{X} &= \pi_l(\pi_{k^\perp}(RX^T))X\\ \dot{R} &= -\pi_{k^\perp}(\pi_l^T(\pi_{k^\perp}(RX^T))R).\end{aligned} \tag{7.7.56}$$

Equations (7.7.56) are Hamiltonian with respect to the canonical structure on $T^*\mathcal{G}$, with corresponding Hamiltonian function

$$H(R,X) = \tfrac{1}{2}\langle \pi_{k^\perp}(RX^T), \pi_{k^\perp}(RX^T)\rangle_T. \tag{7.7.57}$$

From this result, and Lemma 7.7.28, we observe that if $S = \pi_{k^\perp}(R)$ we may rewrite H in (7.7.57) in the form:

$$H = \tfrac{1}{2}\langle \pi_{k^\perp}(SX^T), \pi_{k^\perp}(SX^T)\rangle_T. \tag{7.7.58}$$

Using Lemma 7.7.28 once more, the Hamiltonian equations may be rewritten in the form

$$\begin{aligned} \dot{X} &= \pi_l(\pi_{k^\perp}(SX^T))X \\ \dot{S} &= -\pi_{k^\perp}(\pi_l^T(\pi_{k^\perp}(SX^T))S). \end{aligned} \tag{7.7.59}$$

Now set

$$A = \pi_{k^\perp}(SX^T). \tag{7.7.60}$$

Then the equations (7.7.59) are of the form

$$\begin{aligned} \dot{X} &= \pi_l(\nabla f(A))X \\ \dot{S} &= -\pi_{k^\perp}(\pi_l^T(\nabla f(A))S) \end{aligned}$$

where $f(A) = \frac{1}{2}\mathrm{Tr}(A^2)$.

Now we compute

$$\begin{aligned} \dot{A} &= \pi_{k^\perp}(\dot{S}X^T + S\dot{X}^T) \\ &= \pi_{k^\perp}(-\pi_{k^\perp}(\pi_l^T(\nabla f)S)X^T) \\ &\quad + \pi_{k^\perp}(SX^T\pi_l^T(\nabla f)) \\ &= -\pi_{k^\perp}(\pi_l^T(\nabla f)SX^T) \\ &\quad + \pi_{k^\perp}(SX^T\pi_l^T(\nabla f)) \\ &\quad \text{(by Lemma (7.7.28)} \\ &= \pi_{k^\perp}([A, \pi_l^T(\nabla f)]). \end{aligned}$$

Hence

$$\dot{A} = \pi_{k^\perp}([A, \pi_l^T(\nabla f)]). \tag{7.7.61}$$

But since

$$\nabla f(A) = \pi_l(\nabla f(A)) + \pi_k(\nabla f(A))$$

and $\nabla f(A) \in \mathcal{K}^\perp$ (for any invariant polynomial in A), we have

$$\nabla f(A) = \pi_l^T(\nabla f(A)) - \pi_k(\nabla f(A)).$$

Observing also that $[A, \nabla f(A)] = 0$, we obtain

$$\begin{aligned}
\dot{A} &= \pi_{k^\perp}[A, \pi_k(\nabla f(A))] \\
&= [A, \pi_k(\nabla f(A))] \\
&\quad (\text{since } [\mathcal{K}^\perp, \mathcal{K}] \subset \mathcal{K}^\perp) \\
&= [A, \nabla f(A) - \pi_l \nabla f(A)] \\
&= [\pi_l(\nabla f(A)), A].
\end{aligned}$$

Thus for $f(A) = \frac{1}{2}\mathrm{Tr}(A^2)$, we obtain from (7.7.59) and (7.7.60) the following set of equations which are equivalent to the extremal flow (7.7.56), and represent the (augmented) full Toda flow:

$$\begin{aligned}
\dot{X} &= \pi_l(A)X \qquad (7.7.62)\\
\dot{A} &= [A, \pi_k(A)] = [\pi_l(A), A].
\end{aligned}$$

The equations (7.7.56) and (7.7.59) were originally derived in a completely different fashion in the work of Symes [1980] . We note the special form of the full reduced Toda flow equation (7.7.61)

$$\dot{A} = \pi_{k^\perp}([A, \pi_l^T(\nabla f)]). \qquad (7.7.63)$$

We observe that this evolves naturally in a coadjoint orbit of L. We can see this by considering tangent vectors to an orbit. These are of the form $\mathrm{ad}^*_W A$ for $W \in \mathcal{L}$ and $A \in \mathcal{L}^* \simeq \mathcal{K}^\perp$. Let $A = \langle S, \cdot \rangle_T$ for $S \in \mathcal{K}^\perp$, and let $V \in \mathcal{L}$. Then

$$\begin{aligned}
ad^*_W(A)(V) &= -A(ad_W V) = -A([W, V]) \\
&= -\langle S, [W, V]\rangle_T = -\langle [W^T, S], V\rangle_T \\
&= -\langle [W^T, S], \pi_l V\rangle_T \\
&\quad \text{since } V \in \mathcal{L} \\
&= \langle \pi_{k^\perp}[S, W^T], V\rangle_T.
\end{aligned}$$

Hence indeed the right hand side of (7.7.63) is a tangent vector to a coadjoint orbit. The classical tridiagonal Toda flow lies on a low rank nongeneric orbit of this type.

9

Energy-Based Methods for Stabilization

9.4 Stabilization of a Class of Nonholonomic Systems

Example: Stabilization of Unicycle with Rider

Stabilization of the unicycle with rider is discussed in Zenkov, Bloch, and Marsden [1999], Zenkov, Bloch, and Marsden [2002a]

We now present the dynamical model of a homogeneous disk on a horizontal plane with a mass and pendulum attached. The pendulum is free to move in the plane orthogonal to the disk, while the attached mass stays in the disk's plane. We view this as a simplified model of a rider on a unicycle in which only the sideways motion of the rider (such as the rider's limbs) is modeled, without pedaling control.

Configuration Space. The configuration space for the unicycle with rider as described in Chapter 1, is $Q = S^1 \times S^1 \times S^1 \times SE(2)$, which we parameterize with coordinates $(\theta, \varkappa, \psi, \phi, x, y)$. As in Figure 9.4.1, θ is the tilt of the unicycle itself, $\varkappa$ is that of the limb, and ψ is the angular position of the wheel of the unicycle. The variables (ϕ, x, y), regarded as a point in $SE(2)$, represent the angular orientation and position of the point of contact of the wheel with the ground.

The Symmetry Group. This mechanical system is $SO(2) \times SE(2)$-invariant; the group $SO(2)$ represents the symmetry of the wheel, that is, the symmetry in the ψ variable, while the group $SE(2)$ represents the

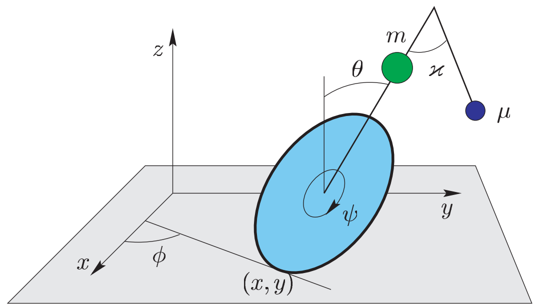

FIGURE 9.4.1. The configuration variables for the unicycle with rider.

Euclidean symmetry of the overall system. The action by the group element (α, β, a, b) on the configuration space is given by

$$
\begin{aligned}
&(\theta, \varkappa, \psi, \phi, x, y) \mapsto \\
&\qquad (\theta, \varkappa, \psi + \alpha, \phi + \beta, x\cos\beta - y\sin\beta + a, x\sin\beta + y\cos\beta + b).
\end{aligned}
$$

Lagrangian. The Lagrangian of this system has the standard form of kinetic minus potential energy:

$$
L = K_{\text{disk}} + \frac{m}{2}v_m^2 + \frac{\mu}{2}v_\mu^2 - U,
$$

where,

$$
\begin{aligned}
K_{\text{disk}} &= \frac{1}{2}\big[A(\dot\theta^2 + \dot\phi^2\cos^2\theta) + B(\dot\phi\sin\theta + \dot\psi)^2\big] \\
&\quad + \frac{M}{2}\big[R^2\dot\theta^2 + 2R(\dot y\cos\phi - \dot x\sin\phi)\dot\theta\cos\theta \\
&\quad + (\dot x - R\dot\phi\sin\theta\cos\phi)^2 + (\dot y - R\dot\phi\sin\theta\sin\phi)^2\big], \\
v_m^2 &= (R+l)^2\dot\theta^2 + 2(R+l)(\dot y\cos\phi - \dot x\sin\phi)\dot\theta\cos\theta \\
&\qquad + \big[\dot x - (R+l)\dot\phi\sin\theta\cos\phi\big]^2 + \big[\dot y - (R+l)\dot\phi\sin\theta\sin\phi\big]^2, \\
v_\mu^2 &= (R+r)^2\dot\theta^2 + \rho^2(\dot\varkappa - \dot\theta)^2 + 2\rho(R+r)(\dot\varkappa - \dot\theta)\dot\theta\cos\dot\varkappa \\
&\quad + 2\big[(R+r)\dot\theta\cos\theta \; + \rho(\dot\varkappa - \dot\theta)\cos(\varkappa - \theta)\big]\big[\dot y\cos\phi - \dot x\sin\phi\big] \\
&\quad + \big[\dot x - \dot\phi\cos\phi\big((R+r)\sin\theta + \rho\sin(\varkappa - \theta)\big)\big]^2 \\
&\quad + \big[\dot y - \dot\phi\sin\phi\big((R+r)\sin\theta + \rho\sin(\varkappa - \theta)\big)\big]^2,
\end{aligned}
$$

and

$$
U = MgR\cos\theta + mg(R+l)\cos\theta + \mu g\left[(R+r)\cos\theta - \rho\cos(\varkappa - \theta)\right].
$$

Constraints. The constraints are given by the standard conditions of rolling without slipping:

$$\dot{x} = -\dot{\psi} R \cos\phi, \qquad \dot{y} = -\dot{\psi} R \sin\phi.$$

Here M is the mass of the disk, R is the radius of the disk, A and B are the principal moments of inertia of the disk, m is the rider mass, r is the rod length, l is the distance from the center of the disk to the mass m, μ is the limb mass, and ρ is the limb length.

Lagrange-d'Alembert Equations. The equations of motion with a control torque u on the pendulum are those derived in the standard way from the Lagrange-d'Alembert principle:

$$\begin{aligned}
&\frac{d}{dt}\frac{\partial L_c}{\partial \dot{\theta}} = \frac{\partial L_c}{\partial \theta},\\
&\frac{d}{dt}\frac{\partial L_c}{\partial \dot{\varkappa}} = \frac{\partial L_c}{\partial \varkappa} + u,\\
&\frac{d}{dt}\frac{\partial L_c}{\partial \dot{\phi}} = \mathcal{A}\cos\theta\,\dot{\theta}\dot{\psi} + \mathcal{B}\cos(\varkappa-\theta)(\dot{\varkappa}-\dot{\theta})\dot{\psi},\\
&\frac{d}{dt}\frac{\partial L_c}{\partial \dot{\psi}} = -\mathcal{A}\cos\theta\,\dot{\theta}\dot{\phi} - \mathcal{B}\cos(\varkappa-\theta)(\dot{\varkappa}-\dot{\theta})\dot{\phi},
\end{aligned}$$

where $L_c(\theta, \varkappa, \dot{\theta}, \dot{\varkappa}, \dot{\psi}, \dot{\phi}) = L(\theta, \varkappa, \phi, \dot{\theta}, \dot{\varkappa}, \dot{\psi}, \dot{\phi}, -\dot{\psi}R\cos\phi, -\dot{\psi}R\sin\phi)$ is the *reduced Lagrangian*, $\mathcal{A} = MR^2 + mR(R+l) + \mu R(R+r)$, and $\mathcal{B} = \mu R \rho$. These equations are supplemented, of course, with the constraints so that one has a well posed initial value problem.

Nonholonomic Momenta and Routhian. The nonholonomic momentum and the constrained Routhian are given by

$$p_1 = \frac{\partial L_c}{\partial \dot{\phi}} = I_{11}\dot{\phi} + I_{12}\dot{\psi}, \qquad p_2 = \frac{\partial L_c}{\partial \dot{\psi}} = I_{21}\dot{\phi} + I_{22}\dot{\psi},$$

and

$$\mathcal{R} = \frac{1}{2}\left(g_{11}\dot{\theta}^2 + 2g_{12}\dot{\theta}\dot{\varkappa} + g_{22}\dot{\varkappa}^2\right) - \frac{1}{2}I^{ab}p_a p_b - U(\theta, \varkappa),$$

respectively. Here,

$$
\begin{aligned}
g_{11} &= MR^2 + m(R+l)^2 + \mu\left[(R+r)^2 - 2(R+r)\rho\cos\varkappa + \rho^2\right] + A,\\
g_{12} &= \mu\left[(R+r)\rho\cos\varkappa - \rho^2\right],\\
g_{22} &= \mu\rho^2,\\
I_{11} &= MR^2\sin^2\theta + m(R+l)^2\sin^2\theta\\
&\quad + \mu\left[(R+r)\sin\theta + \rho\sin(\varkappa-\theta)\right]^2 + A\cos^2\theta + B\sin^2\theta,\\
I_{12} &= MR^2\sin\theta + mR(R+l)\sin\theta\\
&\quad + \mu R\left[(R+r)\sin\theta + \rho\sin(\varkappa-\theta)\right] + B\sin\theta,\\
I_{22} &= MR^2 + mR^2 + \mu R^2 + B.
\end{aligned}
$$

As usual, I^{ab} are the components of the inverse inertia tensor.

Reduced Equations. Using the symmetry of the system, the variables x, y, ϕ, ψ can be eliminated by taking the quotient by the action of the group $SO(2)\times SE(2)$. Carrying this out, the resulting reduced equations of motion may be written in terms of the Routhian as

$$\frac{d}{dt}\frac{\partial\mathcal{R}}{\partial\dot\theta} = \nabla_\theta\mathcal{R}, \tag{9.4.9}$$

$$\frac{d}{dt}\frac{\partial\mathcal{R}}{\partial\dot\varkappa} = \nabla_\varkappa\mathcal{R} + u, \tag{9.4.10}$$

$$\frac{dp_1}{dt} = \left[I^{21}p_1 + I^{22}p_2\right]\left[\mathcal{A}\cos\theta\,\dot\theta + \mathcal{B}\cos(\varkappa-\theta)(\dot\varkappa-\dot\theta)\right], \tag{9.4.11}$$

$$\frac{dp_2}{dt} = -\left[I^{11}p_1 + I^{12}p_2\right]\left[\mathcal{A}\cos\theta\,\dot\theta + \mathcal{B}\cos(\varkappa-\theta)(\dot\varkappa-\dot\theta)\right]. \tag{9.4.12}$$

The covariant derivatives in equations (9.4.9) and (9.4.10) are defined by

$$
\begin{aligned}
\nabla_\theta &= \frac{\partial}{\partial\theta} + \Big[\mathcal{A}\cos\theta - \mathcal{B}\cos(\varkappa-\theta)\Big]\\
&\qquad\times\left[\left(I^{21}p_1 + I^{22}p_2\right)\frac{\partial}{\partial p_1} - \left(I^{11}p_1 + I^{12}p_2\right)\frac{\partial}{\partial p_2}\right],\\
\nabla_\varkappa &= \frac{\partial}{\partial\varkappa} + \mathcal{B}\cos(\varkappa-\theta)\left[\left(I^{21}p_1 + I^{22}p_2\right)\frac{\partial}{\partial p_1} - \left(I^{11}p_1 + I^{12}p_2\right)\frac{\partial}{\partial p_2}\right].
\end{aligned}
$$

The first two of these equations describe the tilting motion of the disk-pendulum system, while the second two model the (coupled) wheel dynamics. The full dynamics is governed by equations (9.4.9)–(9.4.12) coupled with the *reconstruction equation* for the group variables x, y, ϕ, ψ. This reconstruction equation is not needed here as it does not affect the evolution of the shape and the momentum variables, and thus is not used in our stabilization analysis. This is because our stabilization is done modulo the group action, which is natural for the problem. See Zenkov, Bloch, and Marsden [1998] for additional information about the formalism we are using here.

Feedback Law. Now we introduce the following form of the *linear feedback control*

$$u = k_1\theta + k_2\varkappa + k_3\dot{\theta} + k_4\dot{\varkappa}.$$

We can then show that we can achieve linear stability of the pendulum system for a suitable range of parameters and this, coupled with the Lyapunov-Malkin theorem, give overall stability of the system. We omit details here and refer to Zenkov, Bloch, and Marsden [2002a] for the details.

Example: Matching and Energy Methods for Unicycle with Rider

We can also carry out stabilization using the matching techniques of Bloch, Leonard, and Marsden [1998], Bloch, Leonard, and Marsden [2000], Auckly, Kapitanski, and White [2000], Hamberg [1999], see Zenkov, Bloch, Leonard, and Marsden [2000].

Structure of the nonholonomic system. For the unicycle with rider, the symmetry group is $SO(2) \times SE(2)$ and falls into the general class of nonholonomic systems of the form:

1. The *shape space* Q/G is a smooth two-dimensional manifold.
2. The curvature of the *nonholonomic connection* equals zero.
3. The *momentum equation* is of the form of a parallel transport equation.

The reduced equation of motion in this case are

$$\frac{d}{dt}\frac{\partial \mathcal{R}}{\partial \dot{r}^1} = \nabla_1 \mathcal{R}, \qquad \frac{d}{dt}\frac{\partial \mathcal{R}}{\partial \dot{r}^2} = \nabla_2 \mathcal{R} + u, \tag{9.4.13}$$

$$\frac{dp_a}{dt} = \mathcal{D}^b_{a\alpha} p_b \dot{r}^\alpha, \qquad a = 1, \ldots, m, \tag{9.4.14}$$

where $\mathcal{R}$ is the *Routhian*, r^α are the shape variables, p_a are the components of the *nonholonomic momentum*, and the covariant derivatives in equation (9.4.13) are defined by

$$\nabla_\alpha = \partial_{r^\alpha} + \mathcal{D}^b_{a\alpha} p_b \partial_{p_a}. \tag{9.4.15}$$

The term u in the shape equation (9.4.13) represents the control input. The full dynamics is governed by equations (9.4.13) and (9.4.14) coupled with the *reconstruction equation* for the group variables. This reconstruction equation is not needed here as it does not affect the evolution of the shape and the momentum variables, and thus is not used in our stabilization analysis.

The Steady States. The equilibria

$$r = r_0, \qquad p = p^0 \tag{9.4.16}$$

of equations (9.4.13) and (9.4.14) represent the steady state motions of or system. These equilibria are distinguished by the conditions

$$\nabla_1 \mathcal{R} = 0, \qquad \nabla_2 \mathcal{R} = 0$$

and thus are labeled by the p^0. We assume that equilibria (9.4.16) are unstable in the direction of unactuated shape variable.

Matching and Controlled Lagrangians. Matching in this setting proceeds as follows (see Zenkov, Bloch, Leonard, and Marsden [2000], Bloch, Leonard, and Marsden [2000], Hamberg [1999]).

Consider an underactuated system with the Lagrangian $L : TQ \to \mathbb{R}$. Suppose that the configuration variables split into two groups $(q^1, \ldots, q^m)$ and $(q^{m+1}, \ldots, q^n)$ in such a manner that only the equations corresponding to the second group[1] are affected by the control forces:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i} + u_i, \tag{9.4.17}$$

where u_i represent the control inputs and $u_a = 0$ for $a = 1, \ldots, m$. The uncontrolled system has an unstable equilibrium

$$q = q_0, \tag{9.4.18}$$

which we want to stabilize using the control inputs. The controlled Lagrangian approach requires a new function $\widetilde{L}$, the *controlled Lagrangian*, to be constructed such that the equations

$$\frac{d}{dt}\frac{\partial \widetilde{L}}{\partial \dot{q}^i} = \frac{\partial \widetilde{L}}{\partial q^i}, \tag{9.4.19}$$

are equivalent to (9.4.17). We assume that both L and $\widetilde{L}$ are of the form

$$L = \frac{1}{2} g_{ij} \dot{q}^i \dot{q}^j - U(q), \qquad \widetilde{L} = \frac{1}{2} \widetilde{g}_{ij} \dot{q}^i \dot{q}^j - \widetilde{U}(q),$$

with g_{ij}, $\widetilde{g}_{ij}$, U, and $\widetilde{U}$ representing the kinetic energy metrics and the potential energies of the initial and the controlled Lagrangians.

Following Hamberg, we introduce the following tensors:

$$T^i_{jk} = \widetilde{\Gamma}^i_{jk} - \Gamma^i_{jk}, \qquad h^j_i = g_{ik} \widetilde{g}^{kj}, \qquad \widetilde{h}^j_i = \widetilde{g}_{ik} g^{kj} \tag{9.4.20}$$

[1] One can think of this grouping of generalized coordinates more intrinsically as assuming there is a bundle structure $Q \to \mathcal{Q}$, where the fibers of the bundle represent the control directions.

where Γ^i_{jk} and $\widetilde{\Gamma}^i_{jk}$ are the Christoffel symbols of the metrics g_{ij} and $\widetilde{g}_{ij}$. The indices i, j, and k range from 1 to n, and a summation over repeated indices is understood.

The conditions for equivalence of (9.4.17) and (9.4.19) are called the *matching conditions*. They generalize the conditions in Bloch, Leonard, and Marsden [2000] and are given in the following theorem:

9.4.7 Theorem. (Hamberg) *Equations* (9.4.17) *and* (9.4.19) *are equivalent iff the following two conditions hold for* $a = 1, \ldots, m$*:*

$$g_{al} T^l_{jk} = 0, \qquad h^j_a \frac{\partial \widetilde{U}}{\partial q^j} = \frac{\partial U}{\partial q^a}. \tag{9.4.21}$$

The explicit formulae for the controls are

$$u_\alpha = \frac{\partial U}{\partial q^\alpha} - h^j_\alpha \frac{\partial \widetilde{U}}{\partial q^j} - g_{\alpha l} T^l_{jk} \dot{q}^j \dot{q}^k, \qquad \alpha = m+1, \ldots, n. \tag{9.4.22}$$

Of course, the controls provided by Theorem 9.4.7 cannot accomplish *asymptotic stabilization*, which can be gained by adding *dissipative terms* to u_α, *i.e.* by controls of the form

$$u_\alpha = \frac{\partial U}{\partial q^\alpha} - h^j_\alpha \frac{\partial \widetilde{U}}{\partial q^j} - g_{\alpha l} T^l_{jk} \dot{q}^j \dot{q}^k - d_{\alpha j} \dot{q}^j. \tag{9.4.23}$$

For a two degree of freedom system, one can use the following coefficients $d_{\alpha j}$:

$$d_{21} = d\widetilde{h}^2_1, \qquad d_{22} = d\widetilde{h}^2_2, \qquad d > 0. \tag{9.4.24}$$

Feedback Stabilization. Now we use matching techniques or stabilization of of steady state motions

$$r = r_0, \qquad p = p^0 \tag{9.4.25}$$

that satisfy the following condition:

$$\nabla_\alpha \mathcal{R}_{|p=p^0} = \partial_{r^\alpha} \mathcal{R}_{|p=p^0}. \tag{9.4.26}$$

Controlled Routhian. In order to define the constrained Routhian, we introduce an auxiliary *holonomic* system whose Lagrangian is

$$\mathcal{L}_{p^0} = \mathcal{R}_{|p=p^0}.$$

The steady state motion (9.4.25) is dynamically equivalent to the equilibrium of the corresponding auxiliary system. We then apply the holonomic matching technique and obtain the controlled Lagrangian $\widetilde{\mathcal{L}}_{p^0}$. This involves the new metric $\widetilde{g}_{\alpha\beta}$ and the new potential energy, which we construct in the following form:

$$\frac{1}{2} I^{ab} p_a^0 p_b^0 + \widetilde{U}(r).$$

Next, we "unfreeze" the nonholonomic momentum in the above formula. This gives us the *controlled amended potential*

$$\widetilde{U}_a = \frac{1}{2} I^{ab} p_a p_b + \widetilde{U}(r).$$

Finally, we define the controlled Routhian

$$\widetilde{\mathcal{R}} = \frac{1}{2} \widetilde{g}_{\alpha\beta} \dot{r}^\alpha \dot{r}^\beta - \widetilde{U}_a,$$

the *controlled covariant derivatives*

$$\widetilde{\nabla}_\alpha = \frac{\partial}{\partial r^\alpha} + \widetilde{\mathcal{D}}^b_{a\alpha} p_b \frac{\partial}{\partial p_a},$$

and introduce the equations

$$\frac{d}{dt} \frac{\partial \widetilde{\mathcal{R}}}{\partial \dot{r}^\alpha} = \widetilde{\nabla}_\alpha \widetilde{\mathcal{R}}, \qquad \alpha = 1, 2. \tag{9.4.27}$$

The controlled energy corresponding to this controlled Routhian can be chosen, with appropriate choice of gains, to be positive definite at equilibrium (9.4.25).

Nonholonomic Matching. One can then show:

9.4.8 Theorem. *The equations* (9.4.27) *associated with the controlled Routhian* $\widetilde{\mathcal{R}}$ *and the controlled covariant derivatives* $\widetilde{\nabla}$ *coupled with the momentum equations* (9.4.14) *are equivalent to the original equations* (9.4.13) *and* (9.4.14) *iff the following matching conditions hold:*

$$g_{1\gamma} T^\gamma_{\alpha\beta} = 0, \quad h_1^\alpha \frac{\partial \widetilde{V}}{\partial r^\alpha} = \frac{\partial V}{\partial r^1}, \quad h_1^\alpha \left[\frac{1}{2} \frac{\partial I^{bc}}{\partial r^\alpha} + \widetilde{\mathcal{D}}^c_{a\alpha} I^{ab} \right] = \frac{1}{2} \frac{\partial I^{bc}}{\partial r^1} + \mathcal{D}^c_{a1} I^{ab}. \tag{9.4.28}$$

The control u is given by

$$\begin{aligned} u = {} & \frac{\partial U}{\partial q^2} - h_2^\alpha \frac{\partial \widetilde{U}}{\partial q^\alpha} - g_{2\gamma} T^\gamma_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta \\ & + \left(\frac{1}{2} \frac{\partial I^{bc}}{\partial r^2} + \mathcal{D}^c_{a2} I^{ab} - h_2^\alpha \left[\frac{1}{2} \frac{\partial I^{bc}}{\partial r^\alpha} + \widetilde{\mathcal{D}}^c_{a\alpha} I^{ab} \right] \right) p_b p_c. \end{aligned} \tag{9.4.29}$$

One can rewrite the control using the controlled covariant derivatives as

$$u = \nabla_2 U - g_{2\beta} \widetilde{g}^{\alpha\beta} \widetilde{\nabla}_\alpha \widetilde{U} - g_{2\gamma} T^\gamma_{\alpha\beta} \dot{r}^\alpha \dot{r}^\beta. \tag{9.4.30}$$

Nonholonomic Stabilization. In order to stabilize the steady state motion (9.4.25), we will use the results of Zenkov, Bloch, and Marsden [1998] and in particular, the Lyapunov-Malkin theorem. First, we choose the values of gains that produce the controlled Routhian with positive definite at (9.4.25) controlled energy.

One can check that the linearized equations of motion have four pure imaginary and m zero eigenvalues. For the Lyapunov-Malkin theorem to be used, we need to move all the nonzero eigenvalues to the left half plane. We accomplish that by adding the dissipative control terms (9.4.24) to (9.4.30):

$$u = \nabla_2 U - g_{2\beta}\widetilde{g}^{\alpha\beta}\widetilde{\nabla}_\alpha \widetilde{U} - g_{2\gamma}T^\gamma_{\alpha\beta}\dot{r}^\alpha \dot{r}^\beta - d\widetilde{h}^2_\alpha \dot{r}^\alpha. \tag{9.4.31}$$

A direct computation shows that all nonzero eigenvalues of the linearized equation, after the dissipative terms were added, are forced to the left half plane. By the Lyapunov-Malkin theorem, the constructed control stabilizes the steady state motion (9.4.25) along with nearby steady states.

This stabilizes the pendulum system and use of the Lyapunov-Malkin theorem produces stability for the full system. For the details see Zenkov, Bloch, Leonard, and Marsden [2000].

References

Abraham, R. and J. E. Marsden [1978], *Foundations of Mechanics*, Addison-Wesley. Reprinted by Perseus Press, 1995.

Abraham, R., J. E. Marsden, and T. S. Ratiu [1988], *Manifolds, Tensor Analysis, and Applications*, Applied Mathematical Sciences. **75**, Springer-Verlag.

Aeyels, D. [1985], Stabilization by smooth feedback of the angular velocity of a rigid body, *Systems Control Lett.* **6**, 59–63.

Aeyels, D. [1989], Stabilizability and asymptotic stabilizability of the angular velocity of a rigid body, in *New trends in nonlinear control theory (Nantes, 1988)*, Springer, Berlin., 243–253.

Agrachev, A. A. and D. Liberzon [2001], Lie-algebraic stability criteria for switched systems, *SIAM J. Control Optim.* **40**, 253–269 (electronic).

Agrachev, A. A. and A. V. Sarychev [1996], Abnormal sub-Riemannian geodesics: Morse index and rigidity, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13**, 635–690.

Agrachev, A. A. and A. V. Sarychev [1998], On abnormal extremals for Lagrange variational problems, *J. Math. Systems Estim. Control* **8**, 87–118.

Anderson, I. and T. Duchamp [1980], On the existence of global variational principles, *Amer. J. Math.* **102**, 781–862.

Antman, S. S. [1998], The simple pendulum is not so simple, *SIAM Rev* **40**. No. 4. 927–930.

Antman, S. S. and J. E. Osborn [1979], The principle of virtual work and integral laws of motion, *Arch. Rat. Mech. An.* **69**, 231–262.

Appel, P. [1900], Sur l'intégration des équations du mouvement d'un corps pesant de révolution roulant par une arête circulaire sur un plan horizontal; cas parficulier du cerceau, *Rendiconti del circolo matematico di Palermo*, **14**, 1–6.

Appell, P. [1911], Sur lews liaisons experimées par des relations non linéaires entre les vitesses, C. R. Acad. Sc. Paris **152**, 1197–1199.

Arms, J. M. [1981], The structure of the solution set for the Yang–Mills equations, *Math. Proc. Camb. Philos. Soc.***90**, 361–372.

Arms, J. M., J. E. Marsden, and V. Moncrief [1981], Symmetry and bifurcations of momentum mappings, *Comm. Math. Phys.*, **78**, 455–478.

Arms, J. M., J. E. Marsden, and V. Moncrief [1982], The structure of the space solutions of Einstein's equations: II Several Killing fields and the Einstein–Yang–Mills equations, *Ann. of Phys.*, **144**, 81–106.

Arnold, V. I. [1966a], Sur la géométrie differentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits, *Ann. Inst. Fourier, Grenoble*, **16**, 319–361.

Arnold, V. I. [1966b], On an a priori estimate in the theory of hydrodynamical stability, *Izv. Vyssh. Uchebn. Zaved. Mat. Nauk* **54**, 3–5. English Translation: Amer. Math. Soc. Transl. **79** (1969), 267–269.

Arnold, V. I. [1983], *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag.

Arnold, V. I. [1989], *Mathematical Methods of Classical Mechanics*, Second Edition, Springer-Verlag; First Edition 1978, Second Edition, 1989, Graduate Texts in Math, volume **60**.

Arnold, V. I., V. V. Kozlov, and A. I. Neishtadt [1988], *Dynamical Systems III*, Springer-Verlag, Encyclopedia of Math., **3**.

Astolfi, A. [1996], Discontinuous control of nonholonomic systems, *Systems Control Lett.* **27**, 37–45.

Atiyah, M. [1982], Convexity and commuting Hamiltonians, *Bull. London Math. Soc.* **14**, 1–5.

Auckly, D., L. Kapitanski, and W. White [2000], Control of nonlinear underactuated systems, *Comm. Pure Appl.Math.*, **53**, 354–369.

Auslander, L. and R. E. Mackenzie [1977], *Introduction to Differentiable Manifolds*, Dover Publications, New York.

Baillieul, J. [1975], *Some Optimization Problems in Geometric Control Theory*, Ph.D. thesis, Harvard University.

Baillieul, J. [1978], Geometric methods for nonlinear optimal control problems, *J. of Optimization Theory and Applications*, **25**, 519–548.

Baillieul, J. [1987], Equilibrium mechanics of rotating systems, in *Proceedings of the 26th IEEE Conf. Dec. and Control*, 1429–1434.

Baillieul, J. [1990], The behavior of super-articulated mechanisms subject to periodic forcing, in *Analysis of Controlled Dynamical Systems*, (Gauthier and Bride and Bonnard and Kupka, eds.), Birkhäuser.; Proceedings of a Conference held in Lyons, France 3–6 July, 1990.

Baillieul, J. [1993], Stable average motions of mechanical systems subject to periodic forcing, dynamics and control of mechanical systems: in *The Falling Cat and Related Problems* (Michael Enos, ed.) Fields Institute Communications, Am. Math. Soc. Providence. **1**, 1–23.

Baillieul, J. [1995], Energy methods for stability of bilinear systems with oscillatory inputs, in *International Journal of Robust and Nonlinear Control*, Special Issue on the Control of Mechanical Systems, (H. Nijmeijer and A. J. van der Schaft, guest eds.) **5**, 205–381.

Baillieul, J. [1998], The geometry of controlled mechanical systems, in *Mathematical Control Theory*, 322–354. Springer-Verlag, New York.

Baillieul, J. [2000], Kinematic asymmetries and the control of Lagrangian systems with oscillatory inputs, *Proceedings of the IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, 135–143, Pergamon.

Baillieul, J. and J. C. Willems [1999] (eds.) *Mathematical control theory* (Dedicated to Roger Ware Brockett on the occasion of his 60th birthday), Springer-Verlag, N.Y.

Barbashin, E. and N. N. Krasovskii [1952], Stability of motion in the large, *Doklady Mathematics (Translations of Proceedings of Russian Academy of Sciences)* **86**, 3, 453–456.

Barnett, S. [1978], *Introduction to Mathematical Systems Theory*, Oxford University Press, U.K.

Bates, L. [2002], Problems and progress in nonholonomic reduction, *Rep. Math. Phys.* **49**, 143–149. XXXIII Symposium on Mathematical Physics (Torún, 2001).

Bates, L. and R. Cushman [1999], What is a completely integrable nonholonomic dynamical system?, *Rep. Math. Phys.* **44**, 29–35.

Bates, L. and J. Sniatycki [1993], Nonholonomic Reduction, *Reports on Math. Phys.* **32**, 99–115.

Bates, L. [1998], Examples of singular nonholonomic reduction, *Reports in Mathematical Physics* **42**, 231–247.

Bedrosian, N. S. [1992], Approximate feedback linearization: the cart–pole example in *Proceedings 1992 IEEE International Conference on Robotics and Automation*, 1987–1992.

Berry, M. [1984], Quantal phase factors accompanying adiabatic changes, *Proc. Roy. Soc. London A* **392**, 45–57.

Berry, M. [1985], Classical adiabatic angles and quantal adiabatic phase, *J. Phys. A. Math. Gen.* **18**, 15–27.

Berry, M. [1990], Anticipations of the geometric phase, *Physics Today*, 34–40.

Blackall, C. J. [1941], On volume integral invariants of non-holonomic dynamical systems, *Am. J. of Math.* **63**, 155–168.

Blankenstein, G., R. Ortega, and A. van der Schaft [2001], The matching conditions of controlled Lagrangians and interconnection and damping assignment passivity based control *Int. Journal of Control*, (to appear).

Blankenstein, G. and A. J. van der Schaft [2001], Symmetry and reduction in implicit generalized Hamiltonian systems, *Rep. Math. Phys.* **47**, 57–100.

Bliss, G. [1930], The Problem of Lagrange in the Calculus of Variations. *Amer. J. Math.*, 673–744.

Bliss, G. A. [1946], *Lectures on Calculus of Variations*, Univ. of Chicago Press.

Bloch, A. M. [1990], Steepest descent, linear programming and Hamiltonian flows, *Contemp. Math. Amer. Math. Soc.* **114**, 77–88.

Bloch, A. M. [2000], Asymptotic Hamiltonian Dynamics: the Toda lattice, the three-wave interaction and the nonholonomic Chaplygin sleigh, *Physica D*, **141**, 297–315.

Bloch, A. M., R. Brockett, and P. Crouch [1997], Double bracket equations and geodesic flows on symmetric spaces, *Comm. Math. Phys.* **187**, 357–373.

Bloch, A. M., R. W. Brockett, and T. Ratiu [1990], A new formulation of the generalized Toda lattice equations and their fixed-point analysis via the moment map, *Bulletin of the AMS* **23**, 447–456.

Bloch, A. M., R. Brockett, and T. S. Ratiu [1992], Completely integrable gradient flows, *Comm. Math. Phys.* **147**, 57–74.

Bloch, A. M., D. Chang, N. Leonard, and J. E. Marsden [2001], Controlled Lagrangians and the stabilization of mechanical systems II: Potential shaping, *Trans IEEE on Auto. Control* **46**, 1556–1571.

Bloch, A. M., D. Chang, N. Leonard, J. E. Marsden, and C. Woolsey [2000], Asymptotic stabilization of Euler–Poincaré mechanical systems, *Proc. IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, 56–61.

Bloch, A. M. and P. Crouch [1992], On the dynamics and control of nonholonomic systems on Riemannian manifolds, *Proceedings of NOLCOS '92, Bordeaux*, 368–372.

Bloch, A. M. and P. E. Crouch [1993], Nonholonomic and vakonomic control systems on Riemannian manifolds, *Fields Institute Communications* **1**, 25–52.

Bloch, A. M. and P. E. Crouch [1994] Reduction of Euler–Lagrange problems for constrained variational problems and relation with optimal control problems, *The Proceedings of the 33rd IEEE Conference on Decision and Control*, 2584–2590, IEEE (1994).

Bloch, A. M., and P. E. Crouch [1995], Nonholonomic control systems on Riemannian manifolds, *SIAM J. on Control* **37**, 126–148.

Bloch, A. M. and P. Crouch [1996], Optimal control and geodesic flows, *Systems Control Lett.* **28**, no. 2, 65–72.

Bloch, A. M. and P. E. Crouch [1997] Optimal control and the full Toda lattice equations, *Proc 36th IEEE Conf. on Decision and Control*, 1736–1740.

Bloch, A. M. and P. E. Crouch [1998a], Newton's law and integrability of nonholonomic systems, *SIAM J. Control Optim.* **36**, 2020–2039.

Bloch, A. M. and P. E. Crouch [1998b] Optimal control, optimization and analytical mechanics, in *Mathematical Control Theory* (J. Baillieul and J. Willems, eds.), Springer, 268–321.

Bloch, A. M., P. Crouch, J. E. Marsden, and T. S. Ratiu [1998], Discrete rigid body dynamics and optimal control, *Proc. CDC* **37**, 2249–2254.

Bloch, A. M., P. E. Crouch, J. E. Marsden, and T. S. Ratiu [2000], An almost Poisson structure for the generalized rigid body equations, *Proceedings of the IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, 87–92, Pergamon.

Bloch, A. M., P. Crouch, J. E. Marsden, and T. S. Ratiu [2002], The symmetric representation of the rigid body equations and their discretization *Nonlinearity.* **15**, no. 2, 1309–1341.

Bloch, A. M., P. E. Crouch, and T. S. Ratiu [1994] Sub-Riemannian optimal control problems *Fields Institute Communications, AMS.* **3**, 35–48.

Bloch, A. M. and S. V. Drakunov [1994] Stabilization of nonholonomic systems via sliding modes, in *Proceedings of the 33rd IEEE Conference on Decision and Control*, Orlando, Florida, 2961–2963.

Bloch, A. M. and S. V. Drakunov [1995] Tracking in nonholonomic dynamic systems via sliding modes, in *Proceedings of the 34th IEEE Conference on Decision and Control*, New Orleans, LA, 2103–2106.

Bloch, A. M. and S. V. Drakunov [1996], Stabilization and tracking in the nonholonomic integrator via sliding modes, *Systems and Control Letters* **29**, 91–99.

Bloch, A. M. and S. Drakunov [1998], Discontinuous Stabilization of Brockett's Canonical Driftless System, in *Essays in Mathematical Robotics*, *IMA*, 169–183.

Bloch, A. M., S. V. Drakunov, and M. Kinyon [1997], Stabilization of Brockett's generalized canonical driftless system, *Proceedings of the 36th IEEE Conference on Decision and Control*, 4260–4265.

Bloch, A. M., S. V. Drakunov, and M. Kinyon [2000], Stabilization of nonholonomic systems using isospectral flows, *SIAM Journal of Control and Optimization* **38**, 855–874,

Bloch, A. M., H. Flaschka, and T. S. Ratiu [1990], A convexity theorem for isospectral manifolds of Jacobi matrices in a compact Lie algebra, *Duke Math. J.* **61**, 41–65.

Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and G. Sánchez De Alvarez [1992], Stabilization of rigid body dynamics by internal and external torques, *Automatica* **28**, 745–756.

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

Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and T. S. Ratiu [1994], Dissipation Induced Instabilities, *Ann. Inst. H. Poincaré, Analyse Nonlineare* **11**, 37–90.

Bloch, A. M., P. S. Krishnaprasad, J. E. Marsden, and T. S. Ratiu [1996], The Euler–Poincaré equations and double bracket dissipation, *Comm. Math. Phys.* **175**, 1–42.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1997], Stabilization of mechanical systems using controlled Lagrangians, *Proc. of the 36th IEEE Conference on Dec. and Control* 2356–2361.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1998], Matching and stabilization by the method of controlled Lagrangians, *Proc CDC*, **37**, 1446–1451.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1999a], Stabilization of the pendulum on a rotor arm by the method of controlled Lagrangians, *Proc. of the International Conference on Robotics and Automation 1999*, IEEE 500–505.

Bloch, A. M., N. E. Leonard, and J. E. Marsden [1999b], Potential shaping and the method of controlled Lagrangians, *Proc. of the 38th IEEE Conference on Dec. and Control*, 1653–1657.

Bloch, A. M., N. Leonard, and J. E. Marsden [2000], Controlled Lagrangians and the stabilization of mechanical systems I: The first matching theorem, *IEEE Trans. Automat. Control*, **45**, 2253–2270.

Bloch, A. M., N. Leonard, and J. E. Marsden [2001], Controlled Lagrangians and the stabilization of Euler–Poincaré mechanical systems, *Int. J. of Robust and Nonlinear Control* **11**, 191–214.

Bloch, A. M. and J. E. Marsden [1989], Controlling homoclinic orbits, *Theoretical and Computational Fluid Dynamics* **1**, 179–190.

Bloch, A. M., J. E. Marsden, and G. Sánchez De Alvarez [1997], Stabilization of relative equilibria of mechanical systems with symmetry, in *Current and Future Directions in Applied Mathematics*, (M. Alber and B. Hu and J. Rosenthal, eds.), Birkhäuser., 43–64.

Bloch, A. M., M. Reyhanoglu, and H. McClamroch [1992], Control and stabilization of nonholonomic systems, *IEEE Trans. Aut. Control* **37**, 1746–1757.

Bondi, H. [1986], The rigid body dynamics of unidirectional spin, *Proc. Roy. Soc. Lon.* **405**, 265–274.

Boothby, W. M. [1986], *An Introduction to Differentiable Manifolds and Riemannian Geometry, 2nd ed.*, Pure and applied mathematics series **120**, Academic Press, Orlando, FL.

Bourbaki, N. [1971], *Groupes et Algèbres de Lie*, Diffusion C.C.L.S., Paris.

Brockett, R. W. [1970a], Systems theory on group manifolds and coset spaces, *SIAM J. Control* **10**, no. 2, 265–284.

Brockett, R. W. [1970b], *Finite Dimensional Linear Systems*, Wiley.

Brockett, R. W. [1973a], Lie theory and control systems defined on spheres, in *Geometric Methods in System Theory*, *SIAM J. Applied Math.* **10**, no. 2, 213–225.

Brockett, R. W. [1973b], Lie algebras and Lie groups in control theory, in *Geometric Methods in System Theory*, (D. Q. Mayne and R. W. Brockett, eds.), D. Reidel Publ. Co., Dordrecht, the Netherlands., 43–82.

Brockett, R. W. [1976a], Nonlinear systems and differential geometry, *Proc. of the IEEE.* **64**, 61–72.

Brockett, R. W. [1976b], Control theory and analytical mechanics, in *1976 Ames Research Center (NASA) Conference on Geometric Control Theory* (R. Hermann and C. Martin, eds.), Math. Sci. Press, Brookline, Massachussets, Lie Groups: History, Frontiers, and Applications. **VII**, 1–46.

Brockett, R. W. [1976c], Volterra series and geometric control theory, *Automatica—J. IFAC* **12**, 167–176.

Brockett, R. W. [1981], Control theory and singular Riemannian geometry, in *New Directions in Applied Mathematics* (P. J. Hilton and G. S. Young, eds.), Springer-Verlag, 11–27.

Brockett, R. W. [1983], Nonlinear control theory and differential geometry, in *Proc. International Congress of Mathematicians*, 1357–1368.

Brockett, R. W. [1989], On the rectification of vibratory motion, *Sensors and Actuators* **20**, 91–96.

Brockett, R. W. [1991], Dynamical systems that sort lists and solve linear programming problems, *Proc. 27th IEEE Conf. Dec. and Control*, 799–803; see also *Linear Algebra and Its Appl.* **146**, 79–91.

Brockett, R. W. [1994], The double bracket equation as a solution of a variational problem, *Fields Instit. Commun.* **3**, 69–76.

Brockett, R. W. [2000], Beijing lectures on nonlinear control systems, in *Lectures on systems, control, and information (Beijing, 1997)*, *AMS/IP Stud. Adv. Math.* **17**, 1–48.

Brockett, R.W. and L. Dai [1992], Nonholonomic kinematics and the role of elliptic functions in constructive controllability, in *Nonholonomic Motion Planning* (Z. Li and J. F. Canny, eds.), Kluwer, 1–22.

Brockett, R.W. and A. Rahimi [1972], Lie algebras and linear differential equations, in *Ordinary Differential Equations* (L. Weiss, ed.), Academic Press, 379–386.

Bryant, R. and P. Griffiths [1983], Reduction for constrained variational problems and $\int \kappa^2/2\, ds$, *Am. J. of Math.* **108**, 525–570.

Bryant, R. L. and L. Hsu [1993], Rigidity of Integral Curves of Rank 2 distributions, *Inventiones Math.* **114**, 435–461.

Bullo, F. [2000], Stabilization of relative equilibria for underactuated systems on Riemannian manifolds, *Automatica J. IFAC* **36**, 1819–1834.

Bullo, F. [2001], Series expansions for the evolution of mechanical control systems, *SIAM J. Control Optim.* **40**, 166–190.

Bullo, F. [2002] Averaging and vibrational control of mechanical systems, SIAM J. Control Optim. **41**, 542-562.

Bullo, F., N. E. Leonard, and A. D. Lewis [2000], Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups, *IEEE Trans. Automat. Control* **45**, 1437–1454. Mechanics and nonlinear control systems.

Bullo, F., A. D. Lewis, and K. M. Lynch [2002] Controllable kinematic reductions for mechanical systems: concepts, computational tools, and examples *Proceedings of MTNS02*.

Bullo, F. and K. Lynch [2001], Kinematic controllability for decoupled trajectory planning in underactuated mechanical systems, *IEEE Trans. Robotics and Automation* **17**; no. 4, 401–412.

Bullo, F. and R. M. Murray [1999], Tracking for fully actuated mechanical systems: A geometric framework, *Automatica* **35**, 17–34.

Bullo, F. and M. Zefran [2002], On mechanical control systems with nonholonomic constraints and symmetries, *Systems and Control Letters* **45**, 133–142.

Burdick, J., B. Goodwine, and J. P. Ostrowski [1994], The Rattleback Revisited, *Preprint*.

Burke, W. L., *Applied Differential Geometry*, Cambridge University Press.

Byrnes, C. I. and A. Isidori [1988] Local stabilization of minimum-phase nonlinear systems *Systems and Control Letters* **11**, 9–17.

Camarinha, M., F. S. Leite, and P. Crouch [2001], On the geometry of Riemannian cubic polynomials, *Differential Geom. Appl.* **15**, 107–135.

Campion, G., B. D'andréa-Novel, and G. Bastin [1991], Controllability and state feedback stabilizability of nonholonomic mechanical systems. In *Advanced robot control (Grenoble, 1990)*, pages 106–124. Springer, Berlin.

Cantrijn, F. J. M. Cortés, M. de León, and M. de Diego [2002] On the geometry of generlized Chaplygin systems *Math. Proc. Cambridge Phil. Soc.* **132**, 323–351.

Cantrijn, F., M. de León, and M. de Diego [1999], On almost-Poisson structures in nonholonomic mechanics, *Nonlinearity* **12**, 721–737.

Cantrijn, F., M. de León, J. Marrero, and M. de Diego [2000], On almost-Poisson structures in nonholonomic mechanics. II. The time-dependent framework, *Nonlinearity* **13**, 1379–1409.

Cantrijn, F., M. de León, J. Marrero, and M. de Diego [1998], Reduction of nonholonomic mechanical systems with symmetries, *Rep. Math. Phys.* **42**, 25–45.

Canudas De Wit, C. and O. J. Sørdalen [1992], Exponential stabilization of mobile robots with nonholonomic constraints, *MIC—Modeling Identification Control* **13**, 3–14.

Carathéodory, C. [1933] Der Schlitten, *Z. Angew. Math. und Mech.* **13**, 71–76.

Carathéodory, C. [1967] *Calculus of Variations and Partial Differential Equations of the First Order*, Holden Day, San Francisco, CA.

Cardin, F. and M. Favretti [1996], On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints, *J. Geom. and Phys.* **18**, 295–325.

Carr, J. [1981], *Applications of Centre Manifold Theory*, Springer-Verlag.

Cartan, E. [1923], Sur les variétés à connexion affines et la théorie de la relativité géneralisée, *Ann. Ecole Normale, Sup.* **40**, 325–412; See also, **41**, 1–25.

Cartan, E. [1952], Sur la représentation géométrique des systèmes matériels non holonomes, Gauthier-Villars, Paris, Collected Works, Oeuvres Complètes.

Cannas Da Silva, A. and A. Weinstein [1999], *Geometric Models for Noncommutative Alebras*, volume 10 of *Berkeley Mathematics Lecture Notes.* Amer. Math. Soc.

Cendra, H., D. D. Holm, J. E. Marsden, and T. S. Ratiu [1998], Lagrangian Reduction, the Euler–Poincaré Equations and Semidirect Products, *Amer. Math. Soc. Transl.*, **186**, 1–25.

Cendra, H., J. E. Marsden, and T. S. Ratiu [2001a], *Lagrangian reduction by stages*, Memoirs of the Amer. Math. Soc. **152**, Providence, R.I.

Cendra, H., J. E. Marsden, and T. S. Ratiu [2001b], Geometric mechanics, Lagrangian reduction and nonholonomic systems, in *Mathematics Unlimited-2001 and Beyond* (B. Enquist and W. Schmid, eds.), Springer-Verlag, New York., 221–273.

Chang, D., A. M. Bloch, N. Leonard, J. E. Marsden, and C. Woolsey [2002], The equivalence of controlled Lagrangian and controlled Hamiltonian systems, *Control and the Calculus of Variations (special issue dedicated to J. L. Lions).*

Chaplygin, S. A. [1897a], On the motion of a heavy body of revolution on a horizontal plane (Russian), *Physics Section of the Imperial Society of Friends of Physics, Anthropology and Ethnographics, Moscow*, **9**, 10–16; Reproduced in Chaplygin [1954], 413–425.

Chaplygin, S. A. [1897b], On some feasible generalization of the theorem of area, with an application to the problem of rolling spheres (Russian), *Mat. Sbornik* **XX**, 1–32; Reproduced in Chaplygin [1954], 434–454.

Chaplygin, S. A. [1903], On a Rolling Sphere on a Horizontal Plane, *Mat. Sbornik* **XXIV**, 139–168; (Russian) Reproduced in Chaplygin [1949], 72–99 and Chaplygin [1954], 455–471.

Chaplygin, S. A. [1911], On the Theory of the Motion of Nonholonomic Systems. Theorem on the Reducing Factor, *Mat. Sbornik* **XXVIII**, 303–314; (Russian). Reproduced in Chaplygin [1949], 28–38 and Chaplygin [1954], 426–433.

Chaplygin, S. A. [1949], *Analysis of the Dynamics of Nonholonomic Systems*, Moscow, Classical Natural Sciences; (Russian).

Chaplygin, S. A. [1954], *Selected Works on Mechanics and Mathematics*, State Publ. House, Technical-Theoretical Literature, Moscow; (Russian).

Chetaev, N. G. [1959], *The Stability of Motion*, Pergamon Press, New York.

Chow, S. N. and J. K. Hale [1982], *Methods of Bifurcation Theory*, Springer, New York.

Chern, S. J. and J. E. Marsden [1990], A Note on Symmetry and Stability for Fluid Flows, *Geo. Astro. Fluid. Dyn.* **51**, 1–4.

Chernoff, P. R. and J. E. Marsden [1974], *Properties of Infinite Dimensional Hamiltonian systems*, Springer-Verlag, New York, Lecture Notes in Math **124**.

Chorlton, F. [1983] *Textbook of Dynamics*, Ellis Horwood, London.

Chow, W. L. [1939] Über Systemen von linearen partiellen Differentialgleichungen erster Ordnung, *Math. Ann.* **117**, 98–105.

Coddington, E. A. and N. Levinson [1955], *Theory of Ordinary Differential Equations*, McGraw-Hill, New York.

Coron, J. M. [1990] A necessary condition for feedback stabilization, *Systems and Control Letters* **14**, 227–232.

Coron, J. M. [1992], Global asymptotic stabilization for controllable systems without drift, *Mathematics of Controls, Signals and Systems.* **5**, 295–312.

Cortés, J. M. [2002], *Geometric, control and numerical aspects of nonholonomic systems*, Ph.D. thesis, University Carlos III, Madrid (2001) and Springer Lecture Notes in Mathematics.

Cortés, J. M. and M. de León [1999], Reduction and reconstruction of the dynamics of nonholonomic systems, *J. Phys. A: Math. Gen.* **32**, 8615–8645.

Cortés, J. M., M. de León, M. de Diego, and S. Martinez [2001] Mechanical systems subject to generalized constraints, *R. Soc. London. Proc. Ser. A Math. Phys. Eng. Sci* **457**, 651–670.

Cortés. J. and S. Martinez [2000], Optimal control for nonholonomic systems with symmetry, *Proc. 39th IEEE. Conf. Dec. and Cont.*, IEEE, 5216–5218.

Cortés, J M. and and S. Martinez [2002], Configuration controllability of mechanical systems underactuated by one control, *Preprint*, (to appear).

Cortés, J. and S. Martínez [2001], Non-holonomic integrators, *Nonlinearity* **14**, 1365–1392.

Cortés, J. M., S. Martínez, J. P. Ostrowski and H. Zhang [2000], Simple mechanical control systems with constraints and symmetry *SIAM. J. Control. Optim.*, (to appear).

Crabtree, H. [1909], *Spinning Tops and Gyroscopic Motion*, Chelsea.

Crampin, M. [1981], On the differential geometry of the Euler–Lagrange equations, and the inverse problem of Lagrangian dynamics, *J. Phys. A* **14**, 2567–2575.

Crouch, P E. [1977], *Dynamical Realizations of Finite Volterra Series*, Ph.D. thesis, Harvard University.

Crouch, P. E., [1981], Geometric structures in systems theory, *IEEE Proceedings* **128**, 242–252.

Crouch, P. E. [1981], Dynamical realizations of finite Volterra series, *SIAM J. Control Optim.* **19**, 177–202.

Crouch, P. E. [1984], Spacecraft attitude control and stabilization, *IEEE Trans. on Automatic Control*, **29**, 321–333.

Crouch, P. E. and F. Leite [1991], Geometry and the dynamic interpolation problem, *Proc. American Control Conference*, Boston, 1131–1136.

Crouch, P. E. and A. J. van der Schaft [1987], *Variational and Hamiltonian Control Systems*, Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin. **101**.

Crouch, P. E. and F. Silva-Leite [1991a], Geometry and the dynamic interpolation problem, *Proc. A.C.C.*, 1131–1136.

Cushman, R., D. Kemppainen, J. Śniatycki, and L. Bates [1995], Geometry of nonholonomic constraints, *Rep. Math. Phys.* **36**, 275–286.

Cushman, R., J. Hermans, and D. Kemppainen [1996], The rolling disc, in *Nonlinear Dynamical Systems and Chaos (Groningen, 1995)*, Birkhauser, Basel, Boston, MA, Progr. Nonlinear Differential Equations Appl. **19**, 21–60.

Cushman, R. and J. Śniatycki [2002], Nonholonomic reduction for free and proper actions, *Regul. Chaotic Dyn.* **7**, 61–72.

Das, T. and R. Mukherjee [2001], Dynamic analysis of rectilinear motion of a self-propelling disk with unbalanced masses, *Trans. of the ASME* **68**, 58–66.

Davis, M. W. [1987], Some aspherical manifolds, *Duke Math. J.* **5**, 105–139.

Decarlo R., S. Zak, and S. Drakunov [1996], Variable structure and sliding mode control, in *The Control Handbook*, CRC Press, Inc., The Electrical Engineering Handbook Series.

Deift, P., T. Nanda, and C. Tomei [1983], Differential equations for the symmetric eigenvalue problem, *SIAM J. on Numerical Analysis* **20**, 1–22.

De León, M. and M. de Diego [1996], On the geometry of nonholonomic Lagrangian systems, *J. Math. Phys.* **37**, 3389–3414

M. de León and P. Rodgrigues [1989] *Methods of Differential Geometry in Analytical Mechanics*, Mathematics Studies **158**, North Holland, Amsterdam.

Deprit, A. [1983], Elimination of the nodes in problems of N bodies, *Celestial Mech.* **30**, 181–195.

Drakunov, S. and V. I. Utkin [1992] Sliding mode control in dynamic systems, *Int. J. Control* **55**, 1029–1037.

Dubrovin, B. A., A. T. Fomenko and S. P. Novikov [1984], Modern Geometry—Methods and Applications. Part II.The Geometry and Topology of Manifolds, Springer-Verlag, Graduate Texts in Math. **104**.

Ebin, D. G. and J. E. Marsden [1970], Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. Math.*, **92**, 102–163.

Enos, M. J. [1993], *Dynamics and Control of Mechanical Systems*, American Mathematical Society, Providence, RI.; The falling cat and related problems, Papers from the Fields Institute Workshop; Waterloo, Ontario, March 1992.

Erlich, R. and J. Tuszynski [1995], Ball on a rotating turntable: comparison of theory and experiment, *Amer. J. Physics* **63**, 351–359.

Euler, L. [1744], Methodus inveniendi lineas curvas, *Lausannae, Genevae, apud Marcum–Michaelem Bosquet & Socias.*

Favretti, M. [1998], Equivalence of dynamics for nonholonomic systems with transverse constraints, *J. Dynam. Differential Equations* **10**, 511–536 and **29**, 1134–1142.

Faybusovich, L. E. [1988] Explicitly solvable optimal control problems. *Int. J. Control* **48**, no. 6, 235–250.

Fedorov, Y. N. [1985], Integration of a generalized problem on the rolling of a Chaplygin ball., *Geometry, differential equations and mechanics, Moscow*, 151–155.

Fedorov, Y. N. [1989], Two integrable nonholonomic systems in classical dynamics, *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **105**, 38–41.

Fedorov, Y. N. [1994], Generalized Poinsot interpretation of the motion of a multidimensional rigid body, *Trudy Mat. Inst. Steklov* **205**, 200–206.

Fedorov, Y. N. [1999], Systems with an invariant measure on Lie groups. In Simo, C., editor, *Hamiltonian Systems with Three or More Degrees of Freedom*, pages 350–357. Kluwer, NATO ASI Series C, Vol. 533.

Fedorov, Y. N. and V. V. Kozlov [1994], Integrable systems on a sphere with potentials of elastic interaction, *Math. Notes* **54**, 381–386.

Fedorov, Y. N. and V. V. Kozlov [1995], Various aspects of n-dimensional rigid body dynamics, *Amer. Math. Soc. Transl.* **168**, 141–171.

Ferrers [1871], N.M., *Quart. J. of Math.*, **XII**.

Feynman, R. P. [1989], *The Feynman Lectures on Physics*, Addison Wesley, New York.

Filippov A. F., [1988] *Differential equations with discontinuous right–hand sides*, Kluwer Academic Publishers, Boston.

Fischer, A. E., J. E. Marsden, and V. Moncrief [1980], The structure of the space of solutions of Einstein's equations, I: One Killing field, *Ann. Ins. H. Poincaré* **33**, 147–194.

Flaschka, H. [1974], The Toda Lattice, *Phys. Rev. B*, **9**, 1924–1925.

Fufaev, N.A. [1964] On the possibility of realizing a nonholonomic constraint by means of viscous friction forces, *Prikl. Math. Mech* **28**, 513–515.

Gabriel, S. and J. Kajiya [1988], Spline interpolation in curved space, *SIGGRAGH 85*, Course Notes.

Gantmacher, F. R. [1959], *Theory of Matrices.* 2 volumes. Chelsea.

Gauss, C. F. [1829] Über ein neues allgemeines Grundgesatz der Mechanik, *Journal für die Reine und Angewandte Mathematik* **4**, 232–235.

Gelfand, I. M. and S. V. Fomin [1963], *Calculus of Variations.* Prentice-Hall (reprinted by Dover, 2000).

Getz, N. H. and J. E. Marsden [1994], Symmetry and dynamics of the rolling disk, *CPAM Berkeley paper* **630**.

Getz, N. H. and J. E. Marsden [1995], Control for an autonomous bicycle. In *International Conference on Robotics and Automation, IEEE*, Nagoya, Japan.

Getz, N. and J. E. Marsden [1997], Dynamical methods for polar decomposition and inversion of matrices, *Linear Algebra and Its Appl.* **258**, 311–343.

Georgiou, T. T., L. Praly, E. D. Sontag, and A. Teel [1995], Input–output stability, in *The Control Handbook*, CRC Press, Baca Raton.

Giachetta, G. [1992], Jet methods in nonholonomic mechanics, *J. Math. Phys.* **33**, 1652–1665.

Gibbs, J. W. [1879] On the fundamental formulae of dynamics, *Am. J. of Math.* **II**, 49–64.

Godhavn, J.-M. and O. Egeland [1997], A Lyapunov approach to exponential stabilization of nonholonomic systems in power form, *IEEE Trans. Automat. Control* **42**, 1028–1032.

Goldman, W. M. and J. J. Millson [1990], Differential graded Lie algebras and singularities of level sets of momentum mappings, *Comm. Math. Phys.* **131**, 495–515.

Goldstein, H. [1980], *Classical Mechanics*, First Edition 1950, Second Edition 1980, Addison-Wesley.

Gozzi, E. and W. D. Thacker [1987], Classical adiabatic holonomy in a Grassmannian system, *Phys. Rev. D* **35**, 2388–2396.

Greenwood, D. T. [1977] *Classical Dynamics*, Prentice Hall.

Griffiths, P. A. [1983], *Exterior Differential Systems*, Birkhäuser, Boston.

Grossman, R., P. S. Krishnaprasad, and J. E. Marsden [1988], The dynamics of two coupled rigid bodies, in *Dynamical Systems Approaches to Nonlinear Problems in Systems and Circuits* (Salam and Levi, eds.), SIAM., 373–378.

Guckenheimer J. and P. Holmes [1983], *Nonlinear Oscillations, Dynamical Systems and Vector Fields*, Springer-Verlag.

Guichardet, A. [1984], On rotation and vibration motions of molecules, *Ann. Inst. H. Poincaré* **40**, 329–342.

Guillemin, V. and A. Pollack [1974], *Differential Topology*, Prentice-Hall.

Guillemin, V. and S. Sternberg [1978], On the equations of motions of a classic particle in a Yang–Mills field and the principle of general covariance, *Hadronic J.* **1**, 1–32.

Guillemin, V. and S. Sternberg [1984], *Symplectic Techniques in Physics*, Cambridge University Press.

Gurvitz, L. [1992], Averaging approach to nonholonomic motion planning, in *Proc. IEEE Int. Conf. Robotics and Automation*, IEEE Press, Nice, France, 2541–2546.

Hamberg, J. [1999], General Matching Conditions in the Theory of Controlled Lagrangians, in *Proc. of the 38th IEEE Conference on Dec. and Control*, 2519–2523.

Hamel, G. [1904], Die Lagrange–Eulerschen Gleichungen der Mechanik, *Z. für Mathematik u. Physik* **50**, 1–57.

Hamel, G. [1949], *Theoretische Mechanik*, Springer-Verlag.

Hamilton, W. R. [1834], On a general method in dynamics, Part I, *Phil. Trans. Roy. Soc. Lond.*, 247–308.

Hamilton, W. R. [1835], On a general method in dynamics. Part II, *Phil. Trans. Roy. Soc. Lond.*, 95–144.

Hannay, J. [1985], Angle variable holonomy in adiabatic excursion of an integrable Hamiltonian, *J. Phys. A: Math. Gen.*, **18**, 221–230.

Hartman, P. [1982], *Ordinary Differential Equations,* Birkhäuser. (First Edition published by Wiley (1964)).

Helgason, S. [2001] *Differential geometry, Lie groups, and symmetric spaces*, American Mathematical Society, Providence, RI.

Helmke, U. [1994], *Optimization and Dynamical Systems*, Springer-Verlag, New York.

Hermann, R. [1962], Some differential geometric aspects of the Lagrange variational problem, *Indiana Math. J*, 634–673.

Hermann, R. [1963] On the accesibility problem in control theory, in *Nonlinear Differential Equations and Nonlinear Mechanics*, J. P. LaSalle and S. Lefshetz, eds., Academy Press, New York.

Hermann, R. [1977], *Differential Geometry and the Calculus of Variations*, Second Edition, Math. Sci. Press, Brookline, Mass., USA, Interdisciplinary Mathematics.

Hermann, R. and A. J. Krener [1977] Nonlinear controllability and observability *IEEE Trans. Aut. Control* **22**, 728–740.

Hermans, J. [1995], A symmetric sphere rolling on a surface, *Nonlinearity*, **8**, 1–23.

Hermes, H. [1980] On the Synthesis of a Stabilizing Feedback Control via Lie Algebraic Methods, *SIAM Journal on Control and Optimization* **18**, 352–361.

Hespanha, J. P. [1996] Stabilization of the non-holonomic Integrator via logic-based switchings, *Proc. 13th World Congress on Automatic Control, E: Nonlinear Systems*, 467–472.

Hertz, H.R. [1894] *Gessamelte Werke*, Band III. *Der Prinzipien der Mechanik in neuem Zusammenhange dargestellt*, Barth, Leipzig, 1894, English translation MacMillan, London, 1899, reprint Dover, NY, 1956.

Holm, D. D., J. E. Marsden, and T. S. Ratiu [1998], The Euler–Poincaré equations and semidirect products with applications to continuum theories, *Adv. in Math.* **137**, 1–81.

Holm, D. D., J. E. Marsden, T. S. Ratiu, and A. Weinstein [1985], Nonlinear stability of fluid and plasma equilibria, *Phys. Rep.* **123**, 1–196.

Husemoller, D. [1994], *Fibre Bundles,* Springer-Verlag, New York, Graduate Texts in Mathematics, vol. **20**.

Iserles, A. [2002], Expansions that grow on trees, *Notices of the Am. Math. Soc.* **49**, 430–444.

Iserles, A., S. P. Nørsett, and A. F. Rasmussen [2001], Time symmetry and high-order Magnus methods, *Appl. Numer. Math.* **39**, 379–401. Special issue: Themes in geometric integration.

Isidori, A. [1995], *Nonlinear Control Systems*, Springer-Verlag, Berlin, Communications and Control Engineering Series.

Iwai, T. [1987], A gauge theory for the quantum planar three-body system, *J. Math. Phys.* **28**, 1315–1326.

Jalnapurkar, S. M. [1994], Modeling of Constrained Systems, available from http://www.cds.caltech.edu/~smj. (A chapter in the Ph.D. thesis, UC Berkeley, May, 1999).

Jalnapurkar, S. M. and J. E. Marsden [1999], Stabilization of relative equilibria II, *Reg. and Chaotic Dyn.*, **3**, 161–179.

Jalnapurkar, S. M. and J. E. Marsden [2000], Stabilization of Relative Equilibria, *IEEE Trans. Automat. Control* **45**, 1483–1491.

Jiang, Z.-P. and H. Nijmeijer [1999], A recursive technique for tracking control of nonholonomic systems in chained form, *IEEE Trans. Automat. Control* **44**, 265–279.

Jovanović, B. [1998] Nonholonomic Geodesic Flows on Lie Groups and the Integrable Suslov Problem on $SO(4)$. *J. Phys. A: Math. Gen.* **31**, 1451–1422.

Jovanović, B. [2001], Geometry and integrability of Euler–Poincaré–Suslov equations, *Nonlinearity* **14**, 1555–1567.

Jurdjevic, V. [1991], Optimal control problems on Lie groups, in *Analysis of Controlled Dynamical Systems*, B. Bonnard, B. Bride, J. P. Gauthier, and I. Kupka, eds., Birkhäuser, Boston.

Jurdjevic, V. [1993], The Geometry of the Plate–Ball Problem, *Arch. Rat. Mech. An.* **124**, 305–328.

Jurdjevic, V. [1997], *Geometric Control Theory*, Cambridge University Press, Studies in Advanced Mathematics. **52**.

Kalaba, R. E. and F. E. Udwadia [1994], Lagrangian mechanics, Gauss' principle, quadratic programming, and generalized inverses: new equations for nonholonomically constrained discrete mechanical systems, *Quart. Appl. Math.* **52**, 229–241.

Kane, T. R. and D. A. Levinson [1980], Formulation of the equations of motions for complex spacecraft, *J. Guidance and Control* **3**, 99–112.

Kane, T. and M. Scher [1969], A dynamical explanation of the falling cat phenomenon, *Int. J. Solids Structures* **5**, 663–670.

Kapitsa, P. L. [1951] Dynamic stability of a pendulum with a vibrating point of suspension, *Ehksp. Teor. Fiz.* **21**, 588–598.

Karapetyan, A. V. [1980], On the problem of steady motions of nonholonomic systems, *J. Appl. Math. Mech.* **44**, 418–426.

Karapetyan, A. V. [1981], On stability of steady state motions of a heavy solid body on an absolutely rough horizontal plane, *J. Appl. Math. Mech.* **45**, 604–608.

Karapetyan, A. V. [1983], Stability of steady motions of systems of a certain type, *Mechanics of Solids* **18**, 41–47.

Karapetyan, A. V. and V. V. Rumyantsev [1990], Stability of conservative and dissipative systems. In Mikhailov, G. K. and V. Z. Parton, editors, *Applied Mechanics: Soviet Reviews*, volume 1. Hemisphere, NY.

Kazhdan, D., B. Kostant, and S. Sternberg [1978], Hamiltonian group actions and dynamical systems of Calogero type, *Comm. Pure Appl. Math.* **31**, 481–508.

Kelly, S. D. and R. M. Murray [1995], Geometric Phases and Robotic Locomotion, *Journal of Robotic Systems* **12**, 417–431.

Khalil, H. K. [1992], *Nonlinear Systems*, Macmillan Publishing Company, New York.

Kharlamov, P. V. [1992] A critique of some mathematical models of mechanical systems with differential constraints, *J. Appl. Math. Mech.* **54** (4), 584–594.

Khennouf, H. and C. Canudas de Wit [1995] *On the construction of stabilizing discontinuous controllers for nonholonomic systems*, *Proceedings of NOLCOS 95*, 741–746.

Kobayashi, S. and K. Nomizu [1963], *Foundations of Differential Geometry*, Wiley.

Koiller, J. [1992], Reduction of Some Classical Nonholonomic Systems with Symmetry, *Arch. Rat. Mech. An.* **118**, 113–148.

Koiller, J., P. R. Rodrigues, and P. Pitanga [2001], Non-holonomic connections following Élie Cartan, *An. Acad. Brasil. Ciênc.* **73**, 165–190.

Kolmanovsky, I., M. Reyhanoglu, and N. H. McClamroch [1994], Discontinuous feedback stabilization of nonholonomic systems in extended power form, in *Proceedings of the Conf. Dec. and Control*, 3469–3474.

Komarov, A. È. [1990], On the control of nonholonomic systems. In *Problems in the mechanics of controllable motion (Russian)*, 59–64. Perm. Gos. Univ.

Koon, W. S. and J. E. Marsden [1997a], Optimal Control for Holonomic and Nonholonomic Mechanical Systems with Symmetry and Lagrangian Reduction, *SIAM J. Control and Optim.* **35**, 901–929.

Koon, W. S. and J. E. Marsden [1997b], The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems, *Reports on Math Phys.* **40**, 21–62.

Koon, W. S. and J. E. Marsden [1998], The Poisson reduction of nonholonomic mechanical systems, *Reports on Math. Phys.* **42**, 101–134.

Korteweg, D. [1899], Über eine ziemlich verbreitete unrichtige Behandlungsweise eines Problemes der rollenden Bewegung und insbesondere über kleine rollende Schwingungen um eine Gleichgewichtslage, *Nieuw Archiefvoor Wiskunde.* **4**, 130–155.

Kozlov, V. V. [1982a] The dynamics of systems with nonintegrable constraints. I. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **3**, 92–100.

Kozlov, V. V. [1982b] The dynamics of systems with nonintegrable constraints. II. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **4**, 70–76.

Kozlov, V. V. [1982c] The dynamics of systems with nonintegrable constraints. III. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **3**, 102–111.

Kozlov, V. V. [1983], Realization of nonintegrable constraints in classical mechanics, *Sov. Phys. Dokl.* **28**, 735–737.

Kozlov, V. V. [1987] The dynamics of systems with nonintegrable constraints. IV. Integral principles. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **5**, 76–83.

Kozlov, V. V. [1988a] The dynamics of systems with nonintegrable constraints. V. The freedom principle and a condition for ideal constraints. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **6**, 51–54.

Kozlov, V. V. [1988b] Invariant measures of the Euler–Poincaré equations on Lie algebras. *Functional Anal. Appl.* **22**, 69–70.

Kozlov, V. V. [1992] The problem of realizing constraints in dynamics, *J. Appl. Math. Mech.* **56** (4), 594–600.

Kozlov, V. V. and N. N. Kolesnikov [1978], On Theorems of Dynamics, *J. Appl. Math. Mech.* **42**, 28–33.

Krasinskiy, A.Ya. [1988a] On the stability and stabilization of the points of equilibrium of nonholonomic system *J.Appl.Maths.Mechs* **52**, 194-202.

Krasinskiy, A.Ya. [1988b] On the stabilization of the steady motions of of nonholonomic mechanical systems *J.Appl.Maths.Mechs* **52**, 902-908.

Krasovskii, N. [1963], *Stability of Motion*. Stanford University Press (originally published 1959).

Krener, A. J. [1974], A generalization of Chow's theorem and the bang-bang theorem to nonlinear control systems *SIAM J. Control* **12**, 43–52.

Krener, A. J. [1977], The high order maximal principle and its application to singular extremals, *SIAM J. Control Optimization* **15**, 256–293.

Krener, A. J. [1999], Feedback linearization. In *Mathematical control theory*, pages 66–98. Springer, New York.

Krishnaprasad, P. S. [1985], Lie–Poisson structures, dual-spin spacecraft and asymptotic stability, *Nonl. Anal. Th. Meth. and Appl.* **9**, 1011–1035.

Krishnaprasad, P. S. [1989], Eulerian many-body problems, *Cont. Math. AMS* **97**, 187–208.

Krishnaprasad, P. S. and J. E. Marsden [1987], Hamiltonian structure and stability for rigid bodies with flexible attachments, *Arch. Rational Mech. Anal.* **98**, 137–158.

Krishnaprasad, P. and D. P. Tsakiris [2001], Oscillations, SE(2)-snakes and motion control: a study of the Roller Racer, *Dyn. Sys., An International J.* **16**, 347–397.

Krishnaprasad, P. S. and R. Yang [1991] Geometric Phases, Anholonomy and Optimal Movement *Proceedings of Conference on Robotics and Automation, 1991*, Sacramento, California, 2185–2189.

Krishnaprasad, P. S., R. Yang and W. Dayawansa [1991], Control problems on principal bundles and nonholonomic mechanics *Proc. 30th IEEE Conf. Decision and Control*, 1133–1138.

Kummer, M. [1981], On the construction of the reduced phase space of a Hamiltonian system with symmetry, *Indiana Univ. Math. J.* **30**, 281–291.

Lagrange, J. L. [1760], Différents Problémes de Dynamique, *Miscell. Taurin*, **II**; Oeuvres, I, Paris.

Lagrange, J. L. [1788], *Méchanique Analytique*, Chez la Veuve Desaint.

Lafferiere, G. and H. J. Sussman [1991], Motion Planning for Completely Nonholonomic Systems Without Drift, in *Proc. of the Conf. on Robotics and Automation 1991, Sacramento*, 1148–1153.

Laiou, M.-C. and A. Astolfi [1999], Discontinuous control of high-order generalized chained systems, *Systems Control Lett.* **37**, 309–322.

Lasalle, J. and S. Lefschetz [1961], *Stability by Liapunov's Direct Method, with Applications.* Academic Press, NY.

Lee, E. B. and L. Markus [1976] *Foundations of Optimal Control Theory*, J. Wiley and Sons, New York.

Leonard, N. E. and P. S. Krishnaprasad [1993], Averaging for attitude control and motion planning, in *Proceedings of the 32nd IEEE Conference on Decision and Control, December, 1993*, 3098–3104.

Leonard, N. E. and P. S. Krishnaprasad [1995], Motion Control of Drift-Free, Left-Invariant Systems on Lie Groups *IEEE Transactions on Automatic Control* **40**, no. 9, 1539–1554.

Leonard, N. E. and J. E. Marsden [1997], Stability and drift of underwater vehicle dynamics: mechanical systems with rigid motion symmetry, *Physica D*, **105**, 130–162.

Lewis, A. [1995] Controllability of Lagrangian and Hamiltonian Systems, Ph.D. thesis, California Institute of Technology.

Lewis, A. [1996], The geometry of the Gibbs–Appell equations and Gauss' principle of least constraint, *Reports on Math. Phys.* **38**, 11–28.

Lewis, A. [1998], Affine connections and distributions with applications to nonholonomic systems, *Rep. Math. Phys* **42**, 135–164.

Lewis, A. [2000], Simple mechanical control systems with constraints, *IEEE Trans. Aut. Control* **45**, 1420–1436.

Lewis, D. [1995], Linearized dynamics of symmetric Lagrangian systems. In *Hamiltonian dynamical systems*, volume 63 of *IMA Math. Appl.*, pages 195–216. Springer, New York.

Lewis, A. and R. M. Murray [1995], Variational principles in constrained systems: theory and experiments, *Int. J. Nonlinear Mech.* **30**, 793–815.

Lewis, A. D. and R. M. Murray [1997], Controllability of simple mechanical control systems, *SIAM J. on Control and Optimization* **35**, 766–790.

Lewis, A. D. and R. M. Murray [1999], Configuration controllability of simple mechanical control systems, *SIAM Rev.* **41**, 555–574.

Lewis, A., J. P. Ostrowski, R. M. Murray, and J. Burdick [1994], Nonholonomic Mechanics and Locomotion: the Snakeboard Example, in *IEEE Intern. Conf. on Robotics and Automation*.

Lewis, D. and J. C. Simo [1990], Nonlinear stability of rotating pseudo-rigid bodies, *Proc. Roy. Soc. Lon. A* **427**, 281–319.

Li, Z. and J. Canny [1990], Motion of two rigid bodies with rolling constraint, *IEEE Transactions on Robotics and Automation* **6**, 62–71.

Li, Z. and J. Canny [1993], *Nonholonomic Motion Planning*, Kluwer, New York.

Li, Z. and R. Montgomery [1988], Dynamics and Optimal Control of a Legged Robot in Flight Phase, *Proceedings of Conference on Robotics and Automation*, Cincinnati, Ohio, 1816–1821.

Lie, S. [1890], *Theorie der Transformationsgruppen, Zweiter Abschnitt*, Teubner, Leipzig.

Libermann, P. and C. M. Marle [1987], *Symplectic Geometry and Analytical Mechanics.* Kluwer Academic Publishers.

Liu, W. [1992], *Averaging Theorems for Highly Oscillatory Differential Equations and the Approximation of General Paths by Admissable Trajectories for Nonholonomic Systems*, Ph.D. thesis, Rutgers University, NJ.

Liu, W. and H. J. Sussman [1995], Shortest paths for sub-Riemannian metrics on rank-two distributions, *Mem. Amer. Math. Soc.* **118**.

Lobry, C. [1972], *Quelques aspects qualitatifs de la théories de la commande*, Docteur es Sciences Mathématiques, L'Universite Scientifique et Médicale de Grenoble.

Lu, J.-H. and T. S. Ratiu [1991], On the nonlinear convexity theorem of Kostant, *Journ. Amer. Math. Soc.* **4**, 349–364.

Lum, K.-Y., and A. M. Bloch [1999], Generalized Serret–Andoyer transformation and applications for the controlled rigid body. *Dynam. Control* **9**, 39–66.

Lyapunov, A. M. [1907], Problème général de la stabilité du mouvement, *Ann. Fac. Sci. Toulouse* **9**, 203–474.

Lyapunov [1992], The general problem of the stability of motion, *Int. J. Control* **55**, 531–773; translated into English by A. T. Fuller. Also published as a book by Taylor & Francis, London.

Maccullagh, J. [1840], On the rotation of a solid body, *Proc. Roy. Irish Acad. Dublin* **2**, 520–525, (see addenda, pages 542–545, and **3**, 370–371).

Magnus, W. [1954], On the exponential solution of differential equations for a linear operator, *Comm. Pure App. Math.* **7**, 649–673.

Malkin, I. G. [1938], On the Stability of Motion in a Sense of Lyapunov, *Mat. Sbornik*, **XXXV**, 47–101 (Russian).

Manakov, S. V. [1976], Note on the integration of Euler's equations of the dynamics of an n-dimensional rigid body, *Funct. Anal. and its Appl.* **10**, 328–329.

Markeev, A. P. [1983], On Dynamics of a Solid on an Absolutely Rough Plane, *J. Appl. Math. Mech.* **47**, 473–478.

Markeev, A. P. [1992], *The Dynamics of a body contiguous to a solid surface*, Nauka, Moscow (Russian).

Marle, C.-M. [1976], Symplectic manifolds, dynamical groups and Hamiltonian mechanics, in *Differential Geometry and Relativity* (M. Cahen and M. Flato, eds.), D. Reidel, Boston, 249–269.

Marle, C.-M. [1996], Kinematic and geometric constraints, servomechanism and control of mechanical systems, *Rend. Sem. Mat. Univ. Pol. Torino*, **54**, 353–364.

Marle, C.-M. [1998] Various approaches to conservative and nonconservative nonholonomic systems, *Reports on Mathematical Physics* **42**, 211–229.

Marsden, J. E. [1981], *Lectures on Geometric Methods in Mathematical Physics*, CBMS Series, SIAM, Philadelphia, PA.

Marsden, J. E. [1992], *Lectures on Mechanics*, Cambridge University Press, London Mathematical Society Lecture Note Series. **174**.

Marsden, J. E. and M. J. Hoffman [1993], *Elementary Classical Analysis*. W. H. Freeman and Co., NY, Second Edition.

Marsden, J. E. and T. J. R. Hughes [1994], *Foundations of Elasticity*, Prentice Hall, 1983, reprinted by Dover, 1994.

Marsden, J. E., P. S. Krishnaprasad, and J. C. Simo (editors) [1989], *Dynamics and Control of Multibody Systems*, Contemporary Mathematics, Am. Math. Soc. **97**.

Marsden, J. E., R. Montgomery, and T. S. Ratiu [1990], *Reduction, symmetry and phases in mechanics*, Memoirs of the Amer. Math. Soc., Providence, RI, vol **436**.

Marsden, J. E. and J. Ostrowski [1998], Symmetries in Motion: Geometric Foundations of Motion Control, *Nonlinear Sci. Today* (http://link.springer-ny.com).

Marsden, J. E., G. W. Patrick, and S. Shkoller [1998], Multisymplectic Geometry, Variational Integrators, and Nonlinear PDEs, *Commun. Math. Phys.* **1998**, 351–395.

Marsden, J. E., S. Pekarsky, and S. Shkoller [1999], Discrete Euler–Poincaré and Lie–Poisson equations, *Nonlinearity* **12**, 1647–1662.

Marsden, J. E. and T. S. Ratiu [1999], *Introduction to Mechanics and Symmetry*, Springer-Verlag, Texts in Applied Mathematics, **17**; First Edition 1994, Second Edition, 1999.

Marsden, J. E. and J. Scheurle [1993a], Lagrangian Reduction and the Double Spherical Pendulum, *ZAMP* **44**, 17–43.

Marsden, J. E. and J. Scheurle [1993b], The Reduced Euler–Lagrange Equations, *Fields Institute Comm.*, **1**, 139–164.

Marsden, J. E. and A. Weinstein [1970], A comparison theorem for Hamiltonian vector fields, *Proc. AMS* **26**, 629–631.

Marsden, J. E. and A. Weinstein [1974], Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.* **5**, 121–130.

Marsden, J. E. and M. West [2001], Discrete mechanics and variational integrators, *Acta Numerica* **10**, 357–514.

Martínez, S., J. Cortés, and M. De León [2000], The geometrical theory of constraints applied to the dynamics of vakonomic mechanical systems: the vakonomic bracket, *J. Math. Phys.* **41**, no. 4, 2090–2120.

Martínez, S., J. Cortés, and M. D. León [2001], Symmetries in vakonomic dynamics: applications to optimal control, *J. Geom. Phys.* **38**, 343–365.

Maupertuis [1740], Lois du Repos *Mem. de L'Acad. Royale des Sciences de Paris*.

M'Closkey, R. and R. Murray [1993] Convergence rates for nonholonomic systems in power form, in *Proceedings of the American Control Conference*, 2967–2972.

McMillan, W. D. [1936], *Dynamics of Rigid Bodies*, Duncan MacMillan Rowles, UK.

Merkin, D. R. [1997], *Introduction to the theory of stability.* Springer-Verlag, New York. Translated from the third (1987) Russian Edition, edited and with an introduction by Fred F. Afagh and Andrei L. Smirnov.

Meyer, K. R. [1973], Symmetries and integrals in mechanics, in *Dynamical Systems* (M. Peixoto, ed.), Academic Press, 259–273.

Mikhailov, G. K. and V. Z. Parton [1990], *Stability and Analytical Mechanics*, volume 1 of *Applied Mechanics, Soviet Reviews*, Hemisphere Publishing Co.

Milnor, J. [1963], *Morse Theory*, Princeton University Press.

Milnor, J. [1965], *Topology from the Differential Viewpoint*, University of Virginia Press.

Mishchenko, A. S. and A. T. Fomenko [1976], On the integration of the Euler equations on semisimple Lie algebras, *Sov. Math. Dokl.* **17**, 1591–1593.

Mishchenko, A. S. and A. T. Fomenko [1978], Generalized Liouville method of integration of Hamiltonian systems, *Funct. Anal. Appl.* **12**, 113–121.

Misner, C., K. Thorne, and J. Wheeler [1973], *Gravitation*, W. H. Freeman, San Fransisco.

Montaldi, J. A., R. M. Roberts, and I. N. Stewart [1988], Periodic solutions near equilibria of symmetric Hamiltonian systems, *Phil. Trans. R. Soc. Lond. A* **325**, 237–293.

Montgomery, R. [1984], Canonical formulations of a particle in a Yang–Mills field, *Lett. Math. Phys.* **8**, 59–67.

Montgomery, R. [1988], The connection whose holonomy is the classical adiabatic angles of Hannay and Berry and its generalization to the non-integrable case, *Comm. Math. Phys.* **120**, 269–294.

Montgomery, R. [1990], Isoholonomic problems and some applications, *Commun. Math. Phys*, **128**, 565–592.

Montgomery, R. [1991a], Optimal control of deformable bodies and its relation to gauge theory, in *The Geometry of Hamiltonian Systems* (T. Ratiu, ed.), Springer-Verlag, 403–438.

Montgomery, R. [1991b], How much does a rigid body rotate? A Berry's phase from the 18th century, *Amer. J. Phys.* **59**, 394–398.

Montgomery, R. [1993], Gauge theory of the falling cat, *Fields Inst. Commun.* **1**, 193–218.

Montgomery, R. [1994], Abnormal minimizers, *SIAM J. Control Optim.* **32**, 1605–1620.

Montgomery, R. [1995], A survey of singular curves in sub-Riemannian geometry, *J. Dynam. Control Systems* **1**, 49–90.

Montgomery, R. [2002], *A Tour of sub-Riemannian Geometries, their Geodesics and Applications.* Amer. Math. Soc. Mathematical Surveys and Monographs, volume 91.

Montgomery, R., J. E. Marsden, and T. S. Ratiu [1984], Gauged Lie–Poisson structures, in *Fluids and plasmas: geometry and dynamics (Boulder, Colo., 1983)*, Amer. Math. Soc., Providence, RI, 101–114.

Morandi, G., C. Ferrario, G. L. Vecchio, G. Marmo, and C. Rubano [1990], The inverse problem in the calculus of variations and the geometry of the tangent bundle, *Phys. Rep.* **188**, 147–284.

Morgansen, K. A. [2001], Controllability and trajectory tracking for classes of cascade form second-order nonholonomic systems, *Proc. of the 40th IEEE Conference on Dec. and Control* 3031–3036.

Morin, P., J.-B. Pomet, and C. Samson [1999], Design of homogeneous time-varying stabilizing control laws for driftless controllable systems via oscillatory approximation of Lie brackets in closed loop, *SIAM J. Control Optim.* **38**, 22–49.

Morin, P. and C. Samson [2000], Control of nonlinear chained systems: from the Routh–Hurwitz stability criterion to time-varying exponential stabilizers, *IEEE Trans. Automat. Control* **45**, 141–146.

Morse, A. S., editor, [1995], *Control using logic-based switching*, in *Trends in Control*, Papers from the workshop held on Block Island, RI, 1995, Springer-Verlag.

Moser, J. [1974], Finitely many mass points on the line under the influence of an exponential potential — an integrable system, *Springer Lecture Notes in Physics*, **38**, 467–497.

Moser, J. [1975], Three integrable Hamiltonian systems connected with isospectral deformations, *Adv. Math.* **16**, 197–220.

Moser, J. [1976], Periodic orbits near an equilibrium and a theorem by Alan Weinstein, *Comm. Pure Appl. Math.* **29**, 724–747; see also the addendum *Comm. Pure Appl. Math.* **31**, 529–530.

Moser, J. [1980] Various aspects of integrable Hamiltonian systems. In Proc. CIME Conference, Bressanone, Italy 1978, *Prog. Math.* **8**, Birkhäuser.

Moser, J. and A. Veselov [1991], Discrete versions of some classical integrable systems and factorization of matrix polynomials, *Commun. Math. Phys* **139**, 217–243.

Murray, R. [1995], Nonlinear control of mechanical systems: a Lagrangian perspective, in *Proc. IFAC Symposium on Nonlinear Control System Design (NOLCOS)*, 378–389.

Murray, R., Zexiang Li, S. Shankar Sastry [1994], *A mathematical introduction to robotic manipulation*, CRC Press, Boca Raton.

Murray, R. M. and S. Sastry [1993], Nonholonomic motion planning: steering using sinusoids, *IEEE Trans on Automatic Control* **38**, 700–716.

Neimark, J. I. and N. A. Fufaev [1966], On stability of stationary motions of holonomic and nonholonomic systems, *J. Appl. Math. Mech.* **30**, 293–300.

Neimark, J. I. and N. A. Fufaev [1972], *Dynamics of Nonholonomic Systems*, Translations of Mathematical Monographs, AMS, **33**.

Neumann, C. [1888], *Leipzig Berichte*, **XL**.

Newton, I. [1687], *Philosophi Naturalis Principia Mathematica*, Londini Societatis Regiae ac Typis, Josephi and Streater.

Nijmeijer, H. and A J. van der Schaft [1990], *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York.

Noakes, L., G. Heinzinger, and B. Paden [1989], Cubic Splines on Curved Spaces, *IMA Journal of Mathematical Control and Information* **6**, 405–473.

Nomizu, K. [1954], Invariant Affine Connections on Homogeneous Spaces, *Amer. J. Math.* **76**, 33–65.

Oh, Y. G. [1987], A stability criterion for Hamiltonian systems with symmetry, *J. Geom. Phys.*, **4**, 163–182.

Oh, Y. G., N. Sreenath, P. S. Krishnaprasad, and J. E. Marsden [1989], The dynamics of coupled planar rigid bodies Part 2: bifurcations, periodic solutions and chaos, *Dynamics and Diff. Eqns* **1**, 269–298.

Olver, P. J. [1988], Darboux' theorem for Hamiltonian differential operators, *J. Diff. Eqn's* **71**, 10–33.

O'Reilly, O. M. [1996], The Dynamics of Rolling Disks and Sliding Disks, *Nonlinear Dynamics* **10**, 287–305.

Ortega, R., A. Loria, P. J. Nicklasson, and H. Sira-Ramirez [1998], *Passivity-based Control of Euler–Lagrange Systems*, Springer Series in Communications and Control Engineering, Springer-Verlag.

Ortega, R., A. J. van der Schaft, B. Maschke, and G. Escobar [1999], Energy-shaping of port-controlled Hamiltonian systems by interconnection, *Proc. 38th IEEE Conf. Decision Control*, IEEE, 1646–1651.

Ostrowski, J. P. [1995], *Geometric Perspectives on the Mechanics and Control of Undulatory Locomotion*. Ph.D. Thesis, California Institute of Technology.

Ostrowski, J. P. [1998], Reduced equations for nonholonomic mechanical systems with dissipative forces, *Rep. Math. Phys.* **42**, 185–209;

Ostrowski, J. P. [1999], Computing reduced equations for robotic systems with constraints and symmetries, *IEEE Trans. on Robotics and Automation* **15**, 111–113.

Ostrowski, J. P. [2000], Steering for a class of nonholonomic dynamic systems, *IEEE Trans. Aut. Control* **45**, 1492–1497.

Ostrowski, J. P. and J. W. Burdick [1996], Gait kinematics for a serpentine robot, in *IEEE Int. Conf. on Robotics and Automation*, 1294–1299.

Ostrowski, J. P. and J. W. Burdick [1997], Controllability for mechanical systems with symmetries and constraints, *Appl. Math. Comput. Sci.* **7**, 305–331.

Ostrowski, J. P., J. W. Burdick, A. D. Lewis, and R. M. Murray [1995], The mechanics of undulatory locomotion: The mixed kinematic and dynamic case, in *IEEE Intern. Conf. on Robotics and Automation*, 1945–1951.

Ostrowski, J., J. P. Desai, and V. Kumar, [1997], Optimal gait selection for nonholonomic locomotion systems, *Proc. IEEE International Conference on Robotics and Automation*, 786–791.

Pars, L. A. [1965], *A Treatise on Analytical Dynamics*, Heinemann, London.

Pascal, M. [1983], Asymptptic solution of the equations of motion for a Celtic stone, *J. Appl. Math. Mech.* **47**, 269–276.

Pascal, M. [1986], The use of the method of averaging to study nonlinear oscillations of the Celtic stone, *J. Appl. Math. Mech.* **50**, 520–522.

Patrick, G. [1989], The dynamics of two coupled rigid bodies in three space, Amer. Math. Soc., Providence, RI, *Contemp. Math.* **97**, 315–336.

Patrick, G. [1992], Relative equilibria in Hamiltonian systems: The dynamic interpretation of nonlinear stability on a reduced phase space, *J. Geom. and Phys.* **9**, 111–119.

Patrick, G. [1995], Relative equilibria of Hamiltonian systems with symmetry: linearization, smoothness and drift, *J. Nonlinear Sci.* **5**, 373–418.

Polushin, I. G., A. L. Fradkov, and D. D. Khill [2000], Passivity and passification of nonlinear systems, *Avtomat. i Telemekh.*, 3–37.

Pontryagin, L. S. [1959], Optimal control processes, *Usp. Mat. Nauk.* **14**.

Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko [1962], Averaging for attitude control and motion panning, in *The Mathematical Theory of Optimal Processes*, Wiley, New York.

Poincaré, H. [1885], Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation, *Acta. Math.* **7**, 259.

Poincaré, H. [1892–1899]*Les Méthodes Nouvelles de la Mécanique Celeste*. 3 volumes. English translation, *New Methods of Celestial Mechanics*, History of Modern Physics and Astronomy **13**, Amer. Inst. Phys., 1993.

Poincaré, H. [1901], Sur une forme nouvelle des équations de la mécanique, *CR Acad. Sci.* **132**, 369–371.

Pomet, J.-B. [1992], Explicit design of time-varying stabilizing control laws for a class of controllable systems without drift, *Systems and Control Letters* **18**, 147–158.

Rabier, P. J. and W. C. Rheinboldt [2000], *Nonholonomic motion of rigid mechanical systems from a DAE viewpoint*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Ratiu, T. [1980], The motion of the free n-dimensional rigid body, *Indiana U. Math. J.* **29**, 609–627.

Reyhanoglu, M., A. M. Bloch, and N. H. McClamroch [1995], Control of nonholonomic systems with extended base space dynamics, *Internat. J. Robust Nonlinear Control* **5**, 325–330.

Reyhanoglu, M., S. Cho, and N. H. McClamroch [2000], Discontinuous feedback control of a special class of underactuated mechanical systems, *Internat. J. Robust Nonlinear Control* **10**, 265–281.

Reyhanoglu, M., A. J. van der Schaft, N. H. McClamroch, and I. Kolmanovsky [1999], Dynamics and control of a class of underactuated mechanical systems, *IEEE Trans. Automat. Control* **44**, 1663–1671.

Riemann, B. [1860], Untersuchungen über die Bewegung eines flüssigen gleich-artigen Ellipsoides, *Abh. d. Königl. Gesell. der Wiss. zu Göttingen* **9**, 3–6.

Rosenberg, R. M. [1977], *Analytical Dynamics of Discrete Systems*, Plenum Press, NY.

Routh, E. J. [1860], *Treatise on the Dynamics of a System of Rigid Bodies*, Macmillan, London.

Routh, E. J. [1877], *Stability of a given state of motion.* Halsted Press, New York. Reprinted in *Stability of Motion* (1975), A. T. Fuller ed.

Ruina, A. [1998], Non-holonomic stability aspects of piecewise nonholonomic systems, *Reports in Mathematical Physics* **42**, 91–100.

Rumiantsev, Valentin V. [1966], On the stability of the steady motions, *J. Appl. Math. Mech.* **30**, 1090–1103.

Rumiantsev, V. [1978] On Hamilton's principle for nonholonomic systems *J. Appl. Math. Mech.* **42** No. 3, 387–399.

Rumiantsev, V. [1979] On the Lagrange and Jacobi principles for nonholonomic systems *J. Appl. Math. Mech.* **43** No. 4, 583–590.

Rumiantsev, V. [1982a] On integral principles for nonholonomic systems *J. Appl. Math. Mech.* **46** No. 1, 1–8.

Rumiantsev, V. [1982b], On stability problem of a top, *Rend. Sem. Mat. Univ. Padova* **68**, 119–128.

Rund, H. [1966], *The Hamiltonian–Jacobi Theory in the Calculus of Variations*, Krieger, New York.

Ryan, E. P. [1990], Discontinuous feedback and universal adaptive stabilization. In *Control of uncertain systems (Bremen, 1989)*, 245–258. Birkhäuser Boston, MA.

Sánchez de Alvarez, G. [1986], *Geometric Methods of Classical Mechanics Applied to Control Theory*, Ph.D. thesis, Berkeley,

Sánchez de Alvarez, G. [1989], Controllability of Poisson control systems with symmetry, *Cont. Math. AMS* **97**, 399–412.

Sánchez De Alvarez, G. [1993], Poisson brackets and dynamics, in *Dynamical systems (Santiago, 1990)*, Longman Sci. Tech., Harlow, 230–249.

Santilli, R. M. [1978], *Foundations of Theoretical Mechanics, I*, Springer, New York.

Sanders, J. A. and F. Verhulst [1985], *Averaging Methods in Nonlinear Dynamical Systems*, Springer-Verlag, New York, *Applied Mathematical Sciences Series* **59**.

San Martin, L. and P. Crouch [1984], Controllability on principal fibre bundles with compact structure group, *Systems and Control Letters* **5** (1), 35–40.

Sattinger, D. H. and D. L. Weaver [1986], *Lie Groups and Lie Algebras in Physics, Geometry and Mechanics*, Springer-Verlag, New York.

Schneider, D. [2000] *Nonholonomic Euler–Poincaré equations and stability in Chaplygin's sphere*, Ph.D. thesis, University of Washington.

Schneider, D. [2002] Nonholonomic Euler–Poincaré equations and stability in Chaplygin's sphere, *Dynamical Systems*, **17**, no. 2, 87–130.

Schutz, B.F. [1980], *Geometrical Method of Mathematical Physics*, Cambridge University Press.

Seto, D. and J. Baillieul [1994], Control problems in super-articulated mechanical systems, *IEEE Transactions on Automatic Control* (H. Nijmeijer and A. J. van der Schaft, guest eds.) **39–12**, 2442–2453.

Shapere, A. and F. Wilczek [1987], Self-propulsion at low Reynolds number, *Phys. Rev. Lett.* **58**, 2051–2054.

Shapere, A. and F. Wilczek [1988], *Geometric Phases in Physics*, World Scientific, 1988.

Shen, J. [2002] *Nonlinear Control of Multibody Systems with Symmetries via Shape Change*, Ph.D. thesis, University of Michigan.

Shen, J., N. H. McClamroch, and A. M. Bloch [2002] Local equilibrium controllability of multibody systems controlled via shape change, *Preprint*.

Shiriaev, A. S. and A. L. Fradkov [2000], Stabilization of invariant sets for nonlinear non-affine systems, *Automatica J. IFAC* **36**, 1709–1715.

Simo, J. C., D. R. Lewis, and J. E. Marsden [1991], Stability of relative equilibria I: The reduced energy–momentum method, *Arch. Rat. Mech. An.* **115**, 15–59.

Simo, J. C., T. A. Posbergh, and J. E. Marsden [1990], Stability of coupled rigid body and geometrically exact rods: block diagonalization and the energy-momentum method, *Physics Reports* **193**, 280–360.

Simo, J. C., T. A. Posbergh, and J. E. Marsden [1991], Stability of relative equilibria II: Three dimensional elasticity, *Arch. Rational Mech. Anal.* **115**, 61–100.

Simon, B. [1983], Holonomy, the quantum adiabatic theorem and Berry's phase, *Phys. Rev. Letters* **51**, 2167–2170.

Sira-Ramirez, H. [1995], On the sliding mode control of differentially flat systems, *Control-Theory and Advanced Technology* **10**, 1093–1113.

Sira-Ramirez, H. and P. Lischinsky-Arenas [1990], Dynamical discontinuous feedback control of nonlinear systems, *IEEE Trans. Automat. Control* **35**, 1373–1378.

Sjamaar, R. and E. Lerman [1991], Stratified symplectic spaces and reduction, *Ann. of Math.* **134**, 375–422.

Smale, S. [1970], Topology and Mechanics, *Inv. Math.* **10**, 305–331, and **11**, pp. 45–64.

Smart, E. H. [1951], *Advanced Dynamics, Volume II*, McMillan and Co., London.

Sniatycki, J. [1998], Non-holonomic Noether theorem and reduction of symmetries, *Rep. Math. Phys* **42**, 5–23.

Śniatycki, J. [2001], Almost Poisson spaces and nonholonomic singular reduction, *Rep. Math. Phys.* **48**, 235–248.

Śniatycki, J. [2002], The momentum equation and the second order differential equation condition, *Rep. Math. Phys.* **49**, 371–394. XXXIII Symposium on Mathematical Physics (Torún, 2001).

Sommerfeld, A. [1952], *Mechanics, Lectures on Theoretical Physics Vol. 1.* First German Edition 1943, Academic Press.

Sontag, E. D. [1990], *Mathematical control theory. Deterministic finite-dimensional systems*, Springer-Verlag, New York, Texts in Applied Mathematics. **6**.

Sontag, E. [1998], Stability and stabilization: discontinuities and the effect of disturbances, *Proc. NATO Advanced Study Institute "Nonlinear Analysis, Differential Equations and Control,"* F. J. Clarke and R. J. Stern. eds. Kluwer, 551–598.

Sontag, E. D. [1999], Stability and stabilization: discontinuities and the effect of disturbances. In *Nonlinear analysis, differential equations and control (Montreal, QC, 1998)*, 551–598, Kluwer Acad. Publ., Dordrecht.

Sontag, E. D. and Y. Wang [1997], Output-to-state stability and detectability of nonlinear systems, *Systems Control Lett.* **29**, 279–290.

Sontag, E. D. and Y. Wang [1996], New characterizations of input-to-state stability, *IEEE Trans. Automat. Control* **41**, 1283–1294.

Sørdalen, O. J. and O. Egeland [1995], Exponential stabilization of nonholonomic chained systems, *IEEE Trans. Automat. Control* **40**, 35–49.

Spivak, M. [1979], *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Inc., Wilmington, Del.

Spong M. W., Underactuated mechanical systems, in *Control Problems in Robotics and Automation*, B. Siciliano and K. P. Valavanis (eds), *Lecture Notes in Control and Information Sciences* **230** Spinger-Verlag. [Presented at the International Workshop on Control Problems in Robotics and Automation: Future Directions Hyatt Regency, San Diego, California, Dec. 1997.]

Spong, M. W. [1998], On feedback linearization of robot manipulators and Riemannian curvature. In *Essays on Mathematical robotics*, IMA Volumes in Mathematics and Its Applications, Vol. **104**, J. Baillieul, S. S. Sastry, and H. J. Sussmann, eds., Springer-Verlag, NY.

Sreenath, N., Y. G. Oh, P. S. Krishnaprasad, and J. E. Marsden [1988], The dynamics of coupled planar rigid bodies. Part 1: Reduction, equilibria and stability, *Dyn. and Stab. of Systems* **3**, 25–49.

Steenrod, N. E. [1951], *The Topology of Fibre Bundles*, Princeton University Press.

Stephenson, A. [1908] On a new type of dynamical stability, *Mem. Proc. Manch. Lit. Phil. Soc.* **52**, 1–10.

Sternberg, S. [1977], Minimal coupling and the symplectic mechanics of a classical particle in the presence of a Yang–Mills field, *Proc. Nat. Acad. Sci.* **74**, 5253–5254.

Stoker, J. J. [1950], *Nonlinear Vibrations in Mechanical and Electrical Systems*, New York; Republished 1992 in Wiley Classics Library Edition.

Strichartz, R. [1983], Sub-Riemannian geometry, *J. Diff. Geom.* **24**, 221–263; see also *J. Diff. Geom.* **30**, (1989), 595–596.

Strichartz, R. S. [1987] The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations, *Journal of Functional Analysis* **72**, 320–345.

Sumbatov, A. S. [1992], Developments of some of Lagrange's ideas in the works of Russian and Soviet mechanicians, *La Mécanique Analytique de Lagrange et son Héritage, Atti della Accademia delle Scienze di Torino, Suppl. 2* **126**, 169–200.

Sussmann, H. J. [1973], Orbits of families of vector fields and integrability of distributions, *Trans. Amer. Math. Soc.* **180**, 171–180.

Sussmann, H. J. [1979] Subanalytic sets and feedback control, *J. Differential Equations* **31**, 31–52.

Sussmann, H. J. [1986], A product expansion for the Chen series. In *Theory and applications of nonlinear control systems (Stockholm, 1985)*, pages 323–335. North-Holland, Amsterdam.

Sussmann, H. J. [1987] A general theorem on local controllability, *SIAM J. on Control and Optim.* **25**, 158–194.

Sussmann, H. J. [1996], A cornucopia of four-dimensional abnormal sub-Riemannian minimizers. In *Sub-Riemannian geometry*, Edited by A. Bellaïche and J.-J. Risler. Progress in Mathematics, Volume 144, 341–364. Birkhäuser, Basel.

Sussmann, H. J. [1998a], An introduction to the coordinate-free maximum principle, in *Geometry of Feedback and Optimal Control* (B. Jackubczyk and W. Respondek, eds.), Monographs Textbooks Pure. Appl. Math. **207** M. Dekker, 463–557.

Sussmann, H. J. [1998b], Geometry and optimal control. In *Mathematical Control Theory*, J. Baillieul & J. C. Willems, eds., Springer-Verlag, New York, 1998, 140–198.

Sussmann, H. J. and V. Jurdjevic [1972], Controllability of nonlinear systems, *J. Diff. Eqns.* **12**, 95–116; see also Control systems on Lie groups, *J. Differential Equations* **12**, 313–329.

Sussmann, H. J. and W. Liu [1991], Limits of highly oscillatory controls and the approximation of general paths by admissable trajectories, *Proc. 30th IEEE Conf. Decision and Control*, 1190–1195.

Sussmann, H. J. and J. C. Willems [1997], 300 years of optimal control: from the Brachystrochrone to the Maximum Principle, *IEEE Control System Magazine*. **17**, 32–44.

Symes, W. W. [1980] Hamiltonian group actions and integrable systems. *Physica D* **1**, 339–376.

Symes, W. W. [1982], The QR algorithm and scattering for the nonperiodic Toda lattice, *Physica D* **4**, 275–280.

Synge, J. L. and B. A. Griffiths [1959] *Principles of Mechanics*, Third Edition, McGraw Hill.

Tai, M. [2001], *Lagrange Reduction of Nonholonomic Systems on Semidirect Products*, M.A. thesis, Univ. of California, Berkeley.

Teel, A. R. [1996], On graphs, conic relations, and input–output stability of nonlinear feedback systems, *IEEE Trans. Automat. Control* **41**, 702–709.

Teel, A., R. Murray, and G. Walsh [1995], Nonholonomic control systems: from steering to stabilization with sinusoids, *International Journal of Control* **62** (4), 849–870.

Terra, G. and M. H. Kobayashi [2002], On classical mechanical systems with constraints, *Preprint*.

Thimm, A. [1981], Integrable Hamiltonian systems on homogeneous spaces, *Ergodic Theory and Dynamical Systems* **1**, 495–517.

Thomson, W (Lord Kelvin) and P. G. Tait [1879], *Treatise on Natural Philosophy*, Cambridge University Press.

Toda, M. [1975], Studies of a non-linear lattice, *Phys. Rep. Phys. Lett.* **8**, 1–125.

Tomei, C. [1984], The topology of isospectral manifolds of diagonal matrices, *Duke Math. J* **51**, 981–996.

Truesdell, C. [1977], *A First Course in Rational Continuum Mechanics*, Academic Press, New York.

Tsakiris, D. P. [1995] *Motion control and planning for nonholonomic kinematic chains*, Ph.D. thesis, Systems Research Institute, University of Maryland.

Tulczyjew, W. M. [1977], The Legendre transformation, *Ann. Inst. Poincaré* **27**, 101–114.

Utkin V. I. [1978], *Sliding Modes and Their Application in Variable Structure Systems.* Moscow: MIR, 1978.

Utkin, V. I. [1992] *Sliding modes in control and optimization*, Springer-Verlag, New York.

Vaisman, I. [1994], *Lectures on the Geometry of Poisson Manifolds*, volume 118 of *Progress in Mathematics.* Birkhäuser, Boston.

van der Schaft, A. J. [1982], Hamiltonian dynamics with external forces and observations, *Mathematical Systems Theory* **15**, 145–168.

van der Schaft, A. J. [1983], *System Theoretic Descriptions of Physical Systems*, doct. dissertation, University of Groningen; also *CWI Tract #3*, CWI, Amsterdam.

van der Schaft, A. J. [1986], Stabilization of Hamiltonian systems, *Nonlinear An., TMA* **10**, 1021–1035.

van der Schaft, A. J. *L_2-Gain and Passivity Techniques in Nonlinear Control*, Springer-Verlag, Communication & Control Engineering Series.

van der Schaft, A. J. and B. M. Maschke [1994], On the Hamiltonian formulation of nonholonomic mechanical systems, *Rep. on Math. Phys.* **34**, 225–233.

Vela, P. A., K. A. Morgansen, and J. Burdick [2002] Second order averaging methods for oscillatory control of underactuated mechanical systems, *Proc. ACC*, 4672–4677.

Vershik, A. M. [1984], Classical and nonclassical dynamics with constraints (Russian), *Novoe Global. Anal.*, 23–48; see also *Springer Lecture Notes in Mathematics*, **1108**, 278–301.

Vershik, A. M. and L. D. Faddeev [1981], Lagrangian mechanics in invariant form, *Sel. Math. Sov.* **4**, 339–350.

Vershik, A. M. and V. Ya. Gershkovich [1988], Nonholonomic problems and the theory of distributions, *Acta Applicandae Mathematicae* **12**, 181–209.

Vershik, A. M. and V. Y. Gershkovich [1994], Nonholonomic dynamical systems, geometry of distributions and variational problems, in *Dynamical Systems VII*, 1–81. Springer-Verlag, New York,

Vierkandt, A. [1892], Über gleitende und rollende Bewegung, *Monatshefte der Math. und Phys.*, **III**, 31–54.

Voss, A [1885], Ueber die Differentialgleichungen der Mechanik, *Mathematische Annalen* **25**, 258–286.

Vranceneau, G. [1936], *Les Espaces Nonholonomes*, Gauthier-Villars, Paris

Walsh, G. C. and L. G. Bushnell [1995], Stabilization of multiple input chained form control systems, *Systems Control Lett.* **25**, 227–234.

Wang, L. S. and P. S. Krishnaprasad [1992], Gyroscopic control and stabilization, *J. Nonl. Sci.* **2**, 367–415.

Walker, G. T. [1896], On a dynamical top, *Quart. J. Pure Appl. Math.* **28**, 175–184.

Warner, F. W. [1983], *Foundations of Differentiable Manifolds and Lie Groups*, Graduate Texts in Methematics. **94**, Springer-Verlag.

Weber, R. W. [1986], Hamiltonian systems with constraints and their meaning in mechanics, *Arch. Rat. Mech. Anal.*, **91**, 309–335.

Weibel, S. [1997], *Applications of Qualitative Methods in the Nonlinear Control of Superarticulated Mechanical Systems*, PhD thesis, Boston University.

Weibel, S. and J. Baillieul [1998a], Averaging and energy methods for robust open-loop control of mechanical systems, in *Essays on Mathematical Robotics*, J. Baillieul, S.S. Sastry, and H.J. Sussmann, Eds. IMA Series of Springer-Verlag, 203–269.

Weibel, S. and J. Baillieul [1998b], Open-loop stabilization of an n–Pendulum, *Int. J. of Control*, **71**, 931–957.

Weibel, S., T. Kaper, and J. Baillieul [1997], Global dynamics of a rapidly forced cart and pendulum, *Nonlinear Dynamics*, **13**, 131–170.

Wei, J. and E. Norman [1964], On global representations of the solution of linear differential equations as a product of exponentials, *Proc. of the Amer. Math. Soc.*, **15**, 327–334.

Weckesser, W. [1997], A ball rolling on a freely spinning turntable *Am. J. of Phys.* **65**, 736–738.

Weinstein, A. [1971], Symplectic manifolds and their Lagrangian submanifolds, *Advances in Mathematics*, **6**, 329–346.

Weinstein, A. [1973], Normal modes for nonlinear Hamiltonian systems, *Inv. Math.* **20**, 47–57.

Weinstein, A. [1978a], Bifurcations and Hamilton's principle, *Math. Zeit.* **159**, 235–248.

Weinstein, A. [1978b], A universal phase space for particles in Yang–Mills fields, *Lett. Math. Phys.*, **2**, 417–420.

Weinstein, A. [1990], Connections of Berry and Hannay type for moving Lagrangian submanifolds, *Adv. in Math.*, **82**, 133–159.

Weltner, K. [1979], Stable circular orbits of freely moving balls on rotating discs *Am. J. Physics* **47**, 984–986.

Whittaker, E. T. [1988], *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Fourth Edition, Cambridge University Press; First Edition 1904, Fourth Edition, 1937, Reprinted by Dover 1944 and Cambridge University Press, 1988.

Willems, J.C. [1979] System theoretic models for the analysis of physical systems, *Ricerche di Automatica* **10**, 71–106.

Willems, J.C. [1986] From time series to linear systems, I and II *Automatica* **22**, 561–580, 675–694.

Wolf, J. A. [1972], *Spaces of Constant Curvature*, Publish or Perish, Inc., Boston, MA.

Yang, R. [1980], Fibre bundles and the physics of the magnetic monopole, in *Proc. Conference in honor of S.S. Chern*, Springer-Verlag, New York.

Yang, R. [1992], *Nonholonomic Geometry, Mechanics and Control*, Ph.D. thesis, University of Maryland.

Yang, R. and P. S. Krishnaprasad [1994], On the geometry and dynamics of floating four-bar linkages, *Dynam. Stability Systems* **9**, 19–45.

Yang, R., P. S. Krishnaprasad, and W. Dayawansa [1993], Chaplygin dynamics and Lagrangian reduction, *Proc. 2nd Int. Cong. on Nonlinear Mechanics*, (W. Z. Chien, Z. H. Guo, and Y. Z. Guo, eds.), Peking University Press, 745–749.

Yang, R., P. S. Krishnaprasad, and W. Dayawansa [1996], Optimal control of a rigid body with two oscillators. In *Mechanics day (Waterloo, ON, 1992)*, 233–260. Amer. Math. Soc., Providence, RI.

Yano, K. and S. Ishihara [1973], *Tangent and cotangent bundles: differential geometry.* Marcel Dekker Inc., New York. Pure and Applied Mathematics, No. 16.

Zabczyk, J. [1989] Some Comments on Stabilizability, *Applied Math. Optimization* **19**, 1–9.

Zenkov, D. V. [1995], The geometry of the Routh problem, *J. Nonlinear Sci.* **5**, 503–519.

Zenkov, D. V. [1998], *Integrability and Stability of Nonholonomic Systems*, Ph.D. Thesis, The Ohio State University.

Zenkov, D. V. and A. M. Bloch [2000], Dynamics of the n-Dimensional Suslov Problem, *Journal of Geometry and Physics* **34**, 121–136

Zenkov, D. V. and A. M. Bloch, [2002] Invariant measures of nonholonomic flows with internal degrees of freedom (to appear).

Zenkov, D., A. M. Bloch, N. E. Leonard, and J. E. Marsden [2000] Matching and stabilization for the unicycle with rider, *Proc. IFAC Workshop on Lagrangian and Hamiltonian Methods in Nonlinear Control*, 187–188.

Zenkov, D. V., A. M. Bloch, and J. E. Marsden [1998], The energy momentum method for the stability of nonholonomic systems, *Dyn. Stab. of Systems* **13**, 123–166.

Zenkov, D. V., A. M. Bloch, and J. E. Marsden [1999], Stabilization of the unicycle with rider, *Proc. of the 38th IEEE Conference on Dec. and Control*, 3470–3471.

Zenkov, D. V., A. M. Bloch, and J. E. Marsden [2002a], The Lyapunov–Malkin theorem and stabilization of the unicycle with rider, *Systems and Control Lett.* **45**, 293–300.

Zenkov, D. V., A. M. Bloch, and J. E. Marsden [2002b], Flat nonholonomic matching, *Proc ACC*, 2812–2817.

Zenkov, D. V. and V. V. Kozlov [1988], Poinsot's geometric representation in the dynamics of a multidimensional rigid body, *Trudy Sem. Vektor. Tenzor. Anal.*, **23**, 202–204.

Zhang, M. and R. M. Hirschorn [1997], Discontinuous feedback stabilization of nonholonomic wheeled mobile robots, *Dynam. Control* **7**, 155–169.