Chapter 5

Optimal Control

This chapter gives a selection of techniques and results in optimal control theory that are optimization problems for mechanical systems, including nonholonomic systems.

We consider a class of nonlinear optimal control problems, which can be called "optimal control problems in mechanics." We deal with control systems whose dynamics can be described by a system of Euler-Lagrange or Hamilton equations. Using the variational structure of the solution of the corresponding boundary-value problems, we reduce the initial optimal control problem to an auxiliary problem of multiobjective programming. This technique makes it possible to apply some consistent numerical approximations of a multiobjective optimization problem to the initial optimal control problem. For solving the auxiliary problem, we propose an implementable numerical algorithm.

5.1 Variational Nonholonomic Problems

Suppose a submanifold of the tangent bundle is given as the zero set of a set of constraints on the bundle. Suppose also that we are given a Lagrangian or, more generally, an objective function that we wish to minimize or maximize. Then we can proceed in the following two ways:

(1) We can consider the conditional variational problem of minimizing a functional subject to the trajectories lying in the given submanifold and obtain the Euler-Lagrange equations via the Lagrange method of appending the constraints to the Lagrangian via Lagrange multipliers.

(2) We can project, via a suitable projection, the vector field of the unconditional problem on the whole tangent bundle at every point to the tangent space of a given submanifold.
The vector fields arising from these two approaches will not, of course, coincide in general, even though both are tangent to the constraint submanifold.
The two approaches were compared earlier where we compared the two types of dynamics for the vertical rolling disk. The first method gives us variational nonholonomic problems, while (real) nonholonomic mechanics are obtained by a procedure of the second type. In fact, this is implemented by the Lagrange–d'Alembert principle, as we have seen in chapters 1. As we saw, nonholonomic mechanics is not variational, since while we allow all possible variations in taking the variations of the Lagrangian, the variations have to lie in the nonintegrable constraint distribution and are thus not independent of one another or reducible to constraints on the configuration variables.

The Lagrange Problem.

Variational nonholonomic problems, on the other hand, are equivalent to the classical Lagrange problem of minimizing a functional over a class of curves with fixed extreme points and satisfying a given set of equalities.
We have the following :
Let Q be a smooth manifold and TQ its tangent bundle with coordinates $(q^i, \dot{q}^i)$. Let $L : TQ \to R$ be a given smooth Lagrangian and let $\Phi : TQ \to R^{n-m}$ be a given smooth function.

5.1.1 Definition. The Lagrange problem is given by

$$min_{q(\cdot)} \int_0^T L(q, \dot{q}) dt \qquad (5.1.1)$$

subject to the fixed endpoint conditions q(0) = 0, q(T) = qT, and subject to the constraints

$$\Phi(q, \dot{q}) = 0.$$

5.1.2 Example (The Falling Cat Problem). The falling cat problem is an abstraction of the problem of how a falling cat should optimally (in some sense) move its body parts so that it achieves a 180◦ reorientation during its fall.

In this case we begin with a Riemannian manifold Q (the configuration space of the problem) with a free and proper isometric action of a Lie group G on Q (the group SO(3) for the falling cat). Let A denote the mechanical connection; that is, it is the principal connection whose horizontal space is the metric orthogonal to the group orbits. The quotient space $Q/G = X$, the shape space, inherits a Riemannian metric from that on Q. Given a curve $c(t)$ in Q, we shall denote the corresponding curve in the shape space X by $r(t)$.

The problem under consideration is as follows:

Isoholonomic Problem (Falling Cat problem). *Fixing two points* $q_1, q_2 \in Q$*, among all curves* $q(t) \in Q, 0 \leq t \leq 1$*, such that* $q(0) = q_0, q(1) = q_1$*, and* $\dot{q}(t) \in hor_{q(t)}$ *(horizontal with respect to the mechanical connection* ***A****), find the curve or curves* $q(t)$ *such that the energy of the shape space curve, namely,*

$$\frac{1}{2}\int_0^1 \|\dot{r}\|^2 \, dt,$$

is minimized.

Local Solution. We can proceed to solve the Lagrange problem locally by forming the modified Lagrangian

$$\Lambda(q, \dot{q}, \lambda) = L(q, \dot{q}) + \lambda \cdot \Phi(q, \dot{q}), \qquad (5.1.2)$$

with $\lambda \in R^{n-m}$. The Euler-Lagrange equations then take the form

$$\frac{d}{dt}\frac{\partial q}{\partial \dot{q}}\,\Lambda(q,\dot{q}\,,\lambda) - \frac{\partial}{\partial q}\Lambda(q,\dot{q}\,,\lambda) \;=\; 0, \tag{5.1.3}$$

$$\Phi(q,\dot{q}) \;=\; 0. \tag{5.1.4}$$

The case we are particularly interested in is the case of classical (linear in the velocity) nonholonomic constraints:

$$\omega_i(q,\dot{q}) = \sum_{k=1}^{n} a_{ik}(q)\dot{q}^k = 0, \qquad i = 1, \dots, n-m. \tag{5.1.5}$$

In the case that these constraints are integrable (equivalent to functions of q only) and L is physical, i.e., it is a holonomic mechanical system, this system will represent physical dynamics. In the nonholonomic case, these equations will not be physical; one needs the Lagrange–d' Alembert principle. The following theorem gives the differential equations for the Lagrange problem.

5.1.3 Theorem. *A solution of the Lagrange problem* Definition 5.1.1 *with constraints of the form* (5.5.5) *satisfies the following equations:*

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}_i}L - \frac{\partial}{\partial q_i}L + \sum_{j=1}^{n-m}\left(\frac{d}{dt}\lambda_j\right)a_{ji} + \sum_{j=1}^{n-m}\lambda_j\left(\dot{a}_{ji} - \sum_{k=1}^{n}\frac{\partial a_{jk}}{\partial q_i}\,\dot{q}_k\right) = 0 \tag{5.1.6}$$

with the constraints

$$\sum_{k=1}^{n} a_{ik}\,\dot{q}^k = 0 \tag{5.1.7}$$

Contrast these equations of motion with the nonholonomic equations of motion with Lagrange multipliers obtained from the Lagrange–d'Alembert principle:

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}_i}L - \frac{\partial}{\partial q_i}L = \sum_{j=1}^{n-m}\lambda_j\,a_{ji} \tag{5.1.8}$$

Observe that if we (formally) set $\lambda_j = 0$ and $\dot{\lambda}_j = \lambda_j$ in the variational nonholonomic equations, we recover the nonholonomic equations of motion. It is precisely the omission of the $\dot{\lambda}_j$ term that destroys the variational nature of the nonholonomic equations.

Examples 5.1.4. Here we recall two examples that will be used to illustrate the theory above: the vertical rolling penny (or unicycle) and the rolling (homogeneous) ball.

A. (Rolling Disk or Unicycle.)

We consider again the vertical disk this time without controls. The variational problems yielded the augmented Lagrangian

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}\, I\, \dot{\theta}^2 + \frac{1}{2}\, J\, \dot{\varphi}^2 + \mu_1\left(\dot{x} - R\dot{\theta}\cos\varphi\right) + \mu_2\left(\dot{y} - R\dot{\theta}\sin\varphi\right),$$

giving the Lagrange equations

$$m\,\ddot{x} + \dot{\mu}_1 = 0,$$
$$m\,\ddot{y} + \dot{\mu}_2 = 0,$$
$$J\ddot{\varphi} + R\mu_1\dot{\theta}\, sin\,\varphi - R\mu_2\,\dot{\theta}\, cos\,\varphi = 0,$$
$$I\ddot{\theta} - R\frac{d}{dt}\,(\mu_1\, cos\,\varphi + \mu_2\, sin\,\varphi) = 0. \qquad (5.1.9)$$

As we saw previousely in chapter 1, we obtain

$$\mu_1 = -mR\dot{\theta}\, cos\,\varphi + A,$$
$$\mu_2 = -mR\dot{\theta}\, sin\,\varphi + B,$$

where A and B are integration constants giving the equations

$$J\ddot{\varphi} = R\dot{\theta}(Asin\,\varphi - Bcos\,\varphi),$$
$$(I + mR^2)\ddot{\theta} = R\dot{\varphi}\,(-Asin\,\varphi + Bcos\,\varphi).$$

Note that we may obtain the nonholonomic equations of motion by setting the constants of integration for the multipliers A and B equal to zero. However, there is not always so simple a relationship between the variational and nonholonomic equations

Moreover, setting $\mu_j = 0$ and $\mu_j = \dot{\mu}_j$ in equations (5.1.9) gives the equations

$$m\ddot{x} = 0,$$
$$m\ddot{y} = 0,$$
$$J\,\ddot{\varphi} = 0,$$
$$I\,\ddot{\theta} = R(\mu_1 \cos\varphi + \mu_2 \sin\varphi),$$

which are precisely the nonholonomic mechanical equations for the vertical rolling disk (1.4.3), as the theory above indicated.

B. (The Rolling Ball)

Here we treat the example of a controlled rolling ball on the plane as a variational nonholonomic problem. We will use the coordinates x, y for the linear horizontal displacement and P $\in$ SO(3) for the angular displacement of the ball. Thus P gives the orientation of the ball with respect to inertial axes $\mathbf{e_1}$, $\mathbf{e_2}$, $\mathbf{e_3}$, where the e_i are the standard basis vectors aligned with the x−, y −, and z − axes, respectively. In particular, P maps the position of a fixed point in the ball measured in the inertial axes to a fixed reference position. This definition gives a right-invariant description of the kinematics expressed in the body frame, which is useful from some points of view.

Let $\boldsymbol{\omega} \in R^3$ denote the angular velocity of the ball with respect to inertial axes. In particular, the ball may spin freely about the z-axis and the z-component of angular momentum is conserved. If J denotes the inertia tensor of the ball with respect to the body axes, then $J = P^T J P$ denotes the inertia tensor of the ball with respect to the inertial axes, and $\boldsymbol{J\omega}$ is the angular momentum of the ball with respect to the inertial axes. The conservation law alluded to above is expressed as

$$e_3^T \boldsymbol{J\omega} = c. \tag{5.1.10}$$

The nonholonomic constraints are expressed as

$$\begin{aligned} e_3^T \boldsymbol{\omega} + \dot{x} = 0, \\ e_3^T \boldsymbol{\omega} - \dot{y} = 0, \end{aligned} \tag{5.1.11}$$

Note that these do not include constraints on the spin about the z-axis, which can be additionally imposed through applied torques.

The kinematics of the rotating ball may be expressed as $\dot{\boldsymbol{P}} = \boldsymbol{S}(\boldsymbol{\nu})\boldsymbol{P}$, where $\boldsymbol{\nu} = \boldsymbol{P\omega}$ is the angular velocity in the body frame and $\boldsymbol{S}(\boldsymbol{\nu})$ is the skew-symmetric matrix satisfying $a \times b = S(b)a$ for all $a, b \in R^3$. Here we will explicitly derive the Euler–Lagrange equations for the variational nonholonomic problem, from which we may write down the mechanical nonholonomic system.

To obtain the variational control system we first write down the Lagrangian in the following form, where m denotes the mass of the ball:

$$\begin{aligned} L &= \frac{1}{2}\nu^T J \nu + \mu_1(\nu^T \boldsymbol{P} e_1 - \dot{y}) + \mu_2(\nu^T \boldsymbol{P} e_2 - \dot{x}) \\ &+ \frac{1}{2} m(\dot{x}^2 + \dot{y}^2) + \mathrm{trac}\, \mathrm{Q}^{\mathrm{T}}\left(\dot{\boldsymbol{P}} - \boldsymbol{S}(\boldsymbol{\nu})\boldsymbol{P}\right). \end{aligned} \tag{5.1.12}$$

Note that we have expressed the constraints (5.1.11) in terms of ν, and we have treated the kinematic equations themselves as constraints, and have therefore introduced an extra Lagrange multiplier in the form of a matrix Q. (The inner product on the space of 3×3 matrices is just the trace form: $\langle Q, P \rangle = trace\ Q^T P$.)

In order to manipulate the Lagrangian (5.1.12) it is convenient to use the identity

$$a^T \boldsymbol{A}b = trace(ba^T\boldsymbol{A}) = \mathrm{trace}(\boldsymbol{A}baT).$$

The forced Euler-Lagrange equations corresponding to this Lagrangian can now be written as

$$\dot{q}^T + Q^T S(\nu) - \mu_1 e_1 \nu^T - \mu_2 e_2 \nu^T = 0, \tag{5.1.13}$$

$$w^T(J\nu + \mu_1 Pe_1 + \mu_2 Pe_2) - traceQ^T S(\omega)P = 0, \qquad \forall\omega \in R^3, \tag{5.1.14}$$

$$m\ddot{x} + \dot{\mu}_2 = u_1,$$

$$m\ddot{Y} + \dot{\mu}_1 = u_2. \tag{5.1.15}$$

Differentiating equation (5.1.14) yields

$$\omega^T(J\dot{\nu} + \dot{\mu}_1 Pe_1 + \dot{\mu}_2 Pe_2 + \mu_1 S(\nu)Pe_1 + \mu_2 S(\nu)Pe_2)$$

$$- trace\, \dot{q}^T S(\omega)P - traceQ^T S(\omega)\, S(\nu)P = 0,$$

and substituting from (5.1.13) gives

$$\omega^T(J\dot{\nu} + \dot{\mu}_1 Pe_1 + \dot{\mu}_2 Pe_2 + \mu_1 S(\nu)Pe_1 + \mu_2 S(\nu)Pe_2)$$

$$+ \mathrm{trace}\ Q^T (S(\nu)S(\omega) - S(\omega)S(\nu))P$$

$$- \mu_1 \nu^T S(\omega)Pe_1 - \mu_2 \nu^T S(\omega)Pe_2 = 0. \tag{5.1.16}$$

But the Jacobi identity for the cross product yields

$$S(\nu)S(\omega) - S(\omega)S(\nu) = S(S(\nu)\omega), \tag{5.1.17}$$

and from (5.1.9) we obtain

$$\mathrm{trace}\ Q^T S(S(\nu)\omega)P = -\omega^T S(\nu)(J\nu + \mu_1 Pe_1 + \mu_2 Pe_2),$$

so (5.1.16) implies the following system of equations describing the variational controlled rolling ball:

$$J\dot{\nu} = S(\nu)J\nu - \dot{\mu}_1 Pe_1 - \dot{\mu}_2 Pe_2 - \mu_1 S(\nu)Pe_1 + \mu_2 S(\nu)Pe_2$$

$$\dot{P} = S(\nu)P,$$

$$m\ddot{x} = -\dot{\mu}_2 + u_1, \qquad e_2^T P^T \nu + \dot{x} = 0,$$

$$m\ddot{y} = -\dot{\mu}_1 + u_2, \qquad e_1^T P^T \nu - \dot{y} = 0, \tag{5.1.18}$$

Following the prescription described above, we can write down the equations describing the nonholonomic controlled rolling ball in the form

$$
\begin{aligned}
J\dot{\nu} &= S(\nu)J\nu + \lambda_1 Pe_1 + \lambda_2 Pe_2, \\
\dot{P} &= S(\nu)P, \\
m\ddot{x} &= -\lambda_2 + u_1, \qquad e_2^T P^T \nu + \dot{x} = 0, \\
m\ddot{y} &= -\lambda_1 + u_2, \qquad e_1^T P^T \nu - \dot{y} = 0,
\end{aligned}
\tag{5.1.19}
$$

Note that equations (5.1.18) and (5.1.19) can be rewritten in terms of the angular velocity **ω**; the variational equations become

$$
\begin{aligned}
\boldsymbol{J}\dot{\boldsymbol{\omega}} &= S(\boldsymbol{\omega})\boldsymbol{J}\boldsymbol{\omega} - \dot{\mu}_1 Pe_1 - \dot{\mu}_2 Pe_2 - \mu_1 S(\omega)Pe_1 + \mu_2 S(\omega)Pe_2, \\
\dot{P} &= PS(\boldsymbol{\omega}), \\
m\ddot{x} &= -\dot{\mu}_2 + u_1, \qquad e_2^T P^T \boldsymbol{\omega} + \dot{x} = 0, \\
m\ddot{y} &= -\dot{\mu}_1 + u_2, \qquad e_1^T P^T \boldsymbol{\omega} - \dot{y} = 0,
\end{aligned}
$$

while the nonholonomic equations are simply obtained using the usual prescription.[1]

5.2 Variational Nonholonomic Systems and Optimal Control

Variational nonholonomic problems (i.e., constrained variational problems) are equivalent to optimal control problems under certain regularity conditions. We outline the simplest relationship.

Consider a modified Lagrangian

$$
\boldsymbol{\Lambda}(q,\dot{q},\lambda) = L(q,\dot{q}) + \lambda \cdot \boldsymbol{\Phi}(q,\dot{q}) \tag{5.2.1}
$$

with Euler-Lagrange equations

$$
\begin{aligned}
\frac{d}{dt}\frac{\partial}{\partial \dot{q}}\boldsymbol{\Lambda}(q,\dot{q},\lambda) - \frac{\partial}{\partial q}\boldsymbol{\Lambda}(q,\dot{q},\lambda) = 0, \\
\boldsymbol{\Phi}(q,\dot{q}) = 0.
\end{aligned}
\tag{5.2.2}
$$

We will rewrite this equation in Hamiltonian form and show that the resulting equations are equivalent to the equations of motion given by the maximum principle for a suitable optimal control problem.

Set

$$P = \frac{\partial}{\partial \dot{q}} \mathbf{\Lambda}(q, \dot{q}, \lambda) \tag{5.2.3}$$

and consider this equation together with the constraints

$$\mathbf{\Phi}(q, \dot{q}) = 0. \tag{5.2.4}$$

We wish to solve (5.6.3) and (5.6.4) for $(\dot{q}, \lambda)$.

Now assume that on an open set U the matrix

$$\begin{bmatrix} \frac{\partial^2}{\partial \dot{q}^2} \mathbf{\Lambda}(q, \dot{q}, \lambda) & \frac{\partial}{\partial \dot{q}} \mathbf{\Phi}(q, \dot{q})^{\mathrm{T}} \\ \frac{\partial}{\partial \dot{q}} \mathbf{\Phi}(q, \dot{q}) & 0 \end{bmatrix} \tag{5.2.5}$$

has full rank. (This generalizes the usual Legendre condition that $\frac{\partial^2}{\partial \dot{q}^2}$ $\mathrm{L}(\mathrm{q}, \dot{\mathrm{q}})$ has full rank.) By the implicit function theorem, we can solve for $\dot{\mathrm{q}}$ and λ :

$$\dot{\mathrm{q}} = \emptyset(q, p),$$

$$\lambda = \psi(q, p). \tag{5.2.6}$$

We now have the following theorem:

Theorem 5.2.1

Under the transformation (5.2.6), *the Euler–Lagrange system* (5.2.2) *is transformed to the Hamiltonian system*

$$\dot{\mathrm{q}} = \frac{\partial}{\partial p} H(q, p),$$

$$\dot{\mathrm{P}} = -\frac{\partial}{\partial q} H(q, p),$$

$$\tag{5.2.7}$$

where

$$H(q,p) = p \cdot \emptyset(q,p) - L(q,\emptyset(q,p)). \tag{5.2.8}$$

Proof. $\boldsymbol{\Phi}(q,\emptyset(q,p)) = 0$ implies

$$\frac{\partial \boldsymbol{\Phi}}{\partial q} + \frac{\partial \boldsymbol{\Phi}}{\partial \dot{q}} \frac{\partial \emptyset}{\partial q} = 0$$

$$\frac{\partial \boldsymbol{\Phi}}{\partial \dot{q}} \frac{\partial \emptyset}{\partial p} = 0$$

Hence, using (5.2.3), we have

$$\frac{\partial H}{\partial p} = \emptyset + \left(p - \frac{\partial L}{\partial \dot{q}}\right) \cdot \frac{\partial \emptyset}{\partial p} = \dot{q} + \lambda \cdot \left(\frac{\partial \boldsymbol{\Phi}}{\partial \dot{q}} \frac{\partial \emptyset}{\partial p}\right) = \dot{q}.$$

Similarly,

$$\frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q} + \left(p - \frac{\partial L}{\partial \dot{q}}\right) . \frac{\partial \emptyset}{\partial q} = -\frac{\partial L}{\partial q} + \lambda . \left(\frac{\partial \boldsymbol{\Phi}}{\partial \dot{q}} \frac{\partial \emptyset}{\partial q}\right)$$

$$= -\left(\frac{\partial L}{\partial q} + \lambda . \frac{\partial \boldsymbol{\Phi}}{\partial q}\right) = -\frac{\partial \boldsymbol{\Lambda}}{\partial q} = -\dot{p}.$$

We now compare this to the optimal control setup.

Definition 5.2.2. *Let the optimal control problem be given by*

$$min_{u(\cdot)} \int_0^T g(q,u)dt \tag{5.2.9}$$

subject to $q(0) = 0, q(T) = q_T$,

$$\dot{q} = f(q,u),$$

where $q \in R^n, u \in R^m$.

Then we have the following:

Theorem 5.2.3. *The Lagrange problem and optimal control problem generate the same (regular) extremal trajectories, provided that:*

(i) $\Phi(q,\dot{q}) = 0$ *if and only if there exists a u such that* $\dot{q} = f(q,u)$.

(ii) $L(q, f(q,u)) = g(q,u)$.

(iii) *The optimal control* u^* *is uniquely determined by the condition*

$$\frac{\partial \hat{H}}{\partial u}(q,p,u^*) = 0, \tag{5.2.10}$$

where

$$\frac{\partial^2 \hat{H}}{\partial u^2}(q,p,u^*)$$

is of full rank and

$$\hat{H}(q,p,u) = \langle p, f(q,u)\rangle - g(q,u) \tag{5.2.11}$$

is the Hamiltonian function given by the maximum principle.

Proof. By (iii) we may use the equation

$$p \cdot \frac{\partial f}{\partial u}(q,u^*) - \frac{\partial g}{\partial u}(q,u^*) = 0$$

to deduce that there exists a function r such that $u^* = r(q,p)$.

The extremal trajectories are now generated by the Hamiltonian

$$\begin{aligned} \bar{H}(q,p) &= \hat{H}\left(q,p,r(x,p)\right) \\ &= p \cdot f\big(q,r(q,p)\big) - g\big(q,r(q,p)\big). \end{aligned} \tag{4.6.12}$$

Then the result follows, and we have

$$\bar{H}(q,p) = H(q,p),$$

$$f(q, r(q,p)) = \emptyset(q,p),$$
$$g(q, r(q,p)) = L(q, \emptyset(q,p)). \quad [10]$$

5.3 Optimal Control of a Homogeneous Ball on a Rotating Plate

Bloch, Krishnaprasad, Marsden and Murray also studies a well-known example, namely the model of a homogeneous ball on a rotating plate and writes down its equations of motion in a form that is suitable for the application of control theory. Fix coordinates in inertial space and let the plane rotate with constant angular velocity about the z-axis. The configuration space of the sphere is Q = $\mathbb{R}^2 \times SO(3)$, parameterized by (x, y, g), g $\in$ SO(3), all measured with respect to the inertial frame. Let $\omega = (\omega_x, \omega_y, \omega_z)$ be the angular velocity vector of the sphere measured also with respect to the inertial frame, let m be the mass of the sphere, mk^2 its inertia about any axis, and let α be its radius.

The Lagrangian of the system is

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}mk^2(\omega_x^{\,2} + \omega_y^{\,2} + \omega_z^{\,2})$$

with the affine nonholonomic constraints

$$\dot{x} - \alpha\omega_y = -\Omega_y$$
$$\dot{y} + \alpha\omega_x = \Omega_x$$

Note that the Lagrangian here is a metric on Q which is bi-invariant on SO(3) as the ball is homogeneous. Note also that $\mathbb{R}^2 \times SO(3)$ is a principal bundle over $\mathbb{R}^2$ with respect to the right SO(3) action on Q given by

$$(x, y, g) \mapsto (x, y, gh)$$

for h $\in$ SO(3). The action is on the right since the symmetry is a material symmetry.

After some computations, it can be shown that the equations of motion are:

$$\omega_x + \frac{1}{\alpha}\dot{y} = \frac{\Omega x}{\alpha}$$

$$\omega_y - \frac{1}{a}\dot{x} = \frac{\Omega y}{a}$$

$$\omega_z = c,$$

(where c is a constant), together with

$$\ddot{x} + \frac{k^2\Omega}{a^2+k^2}\dot{y} = 0$$

$$\ddot{y} - \frac{k^2\Omega}{a^2+k^2}\dot{x} = 0 \qquad (5.3.1)$$

Notice that the first set of three equations has the form

$$\dot{g}\, g^{-1} = A_{loc}(r)\dot{r} + \Gamma_{loc}(r)$$

where

$$A_{loc} = \frac{1}{a}e_1 d_y - \frac{1}{a}e_2 d_x$$

and

$$\Gamma_{loc} = \frac{\Omega}{a}xe_1 + \frac{\Omega}{a}ye_2 + ce_3\ . \qquad (5.3.2)$$

Here, $r^1 = x$, $r^2 = y$ and e_1, e_2, e_3 is the standard basis of so(3). Also, A_{loc} is the expression of nonholonomic connection relative to the (global) trivialization and Γ_{loc} is the expression of the affine piece of the constraints with respect to the same trivialization.

Now we are ready to apply reduced Lagrangian optimization to find the optimal trajectories for a homogeneous ball. Clearly the homogeneous ball on a rotating plate is a simple nonholonomic mechanical system with symmetry as defined earlier, which also has a trivial principal bundle structure (except that the constraint is affine which can be dealt with in the same way). Also we can assume that we have full control over the motion of the center of the ball, i.e., over the shape variables. Now let the cost function be $C(\dot{r}) = \frac{1}{2}[(\dot{x})^2 + (\dot{y}^2)]$ and set $a = 1$ for simplicity, then we can use the method of Lagrange multipliers and Lagrangian reduction to find the necessary conditions for the optimal trajectories of the following optimal control problem:

Plate Ball Problem

Given two points q_0, $q_1 \in \mathbb{R}^2 \times SO(3)$, find the optimal control curves (x(t), y(t)) $\in \mathbb{R}^2$ that steer the system from q_0 to q_1 and minimizes $\int_0^1 \frac{1}{2}[(\dot{x})^2 + (\dot{y})^2]\, dt$, subject to the constraints

$$\dot{g}g^{-1} = -\dot{y}\, e_1 + \dot{x}e_2 + ce_3 + \Omega x e_1 + \Omega y e_2, \qquad (5.3.3)$$

where, again, e_α is the standard basis of so(3) .

Following the reduced Lagrangian optimization method developed in the preceding section, we

define a new Lagrangian $\mathcal{L}$ by

$$\mathcal{L} = \frac{1}{2}[(\dot{x})^2 + (\dot{y})^2] + \lambda_\alpha \xi^\alpha + \lambda_1 \dot{y} - \lambda_2 \dot{x} - \lambda_3 c - \Omega\lambda_1 x - \Omega\lambda_2 y, \qquad (5.3.4)$$

where $\lambda(t) \in so(3)$.

By the preceding Theorem, we know that any the reduced optimal curve (x(t), y(t), $\dot{x}$(t), $\dot{y}$(t), ξ^α (t)) must satisfy the reduced Euler Lagrangian equations. Simple computations show that

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = \dot{x} - \lambda_2 = \rho_1$$

$$\frac{\partial \mathcal{L}}{\partial x} = -\Omega\lambda_1$$

$$\frac{\partial \mathcal{L}}{\partial \dot{y}} = \dot{y} + \lambda_1 = \rho_2$$

$$\frac{\partial \mathcal{L}}{\partial y} = -\Omega\lambda_2$$

$$\frac{\partial \mathcal{L}}{\partial \xi^b} = \lambda_b.$$

Therefore

$$\dot{\rho}_1 = -\Omega\lambda_1$$

$$\dot{\rho}_2 = -\Omega\lambda_2,$$

and

$$\dot{\lambda}_b = C^a_{db}\lambda_a \xi^d,$$

that is:

$$\dot{\lambda}_1 = \lambda_3\xi^2 - \lambda_2\xi^3 = \lambda_3(\rho_1 + \lambda_2 + \Omega_y) - c\lambda_2$$

$$\dot{\lambda}_2 = -\lambda_3\xi^1 + \lambda_1\xi^3 = \lambda_3(\rho_2 - \lambda_1 - \Omega_x) + c\lambda_1$$

$$\dot{\lambda}_3 = \lambda_2\xi^1 - \lambda_1\xi^3 = -(\lambda_3\rho_1 + \lambda_2\rho_2) + \Omega(\lambda_2 x + \lambda_1 y). \qquad (5.3.5)$$

In the special case where c = 0 (no drift) and Ω = 0 (no rotation), we have

$$\dot{\rho}_1 = 0$$

$$\dot{\rho}_2 = 0$$

$$\dot{\lambda}_1 = \lambda_3(\rho_1 + \lambda_2)$$

$$\dot{\lambda}_2 = \lambda_3(\rho_2 - \lambda_1)$$

$$\dot{\lambda}_3 = -(\lambda_1\rho_1 + \lambda_2\rho_2). \qquad (5.3.6)$$

which gives the same result. [2]

5.4 Optimal Control of the Snakeboard

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.

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. Thus the configuration space is Q = SE(2)$\times S^1 \times S^1 \times S^1$. 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.

The Lagrangian for the snakeboard consists only of kinetic energy terms and can be written as

$$L(q, \dot{q}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}J\dot{\theta}^2 + \frac{1}{2}J_0(\dot{\theta} + \dot{\psi})^2 + \frac{1}{2}J_1(\dot{\theta} + \dot{\phi}_1)^2 + \frac{1}{2}J_2(\dot{\theta} + \dot{\phi}_2)^2, \quad (5.4.1)$$

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_i, i = 1, 2, is the inertia corresponding to ϕ_i. The Lagrangian is independent of the configuration of the board and hence it is invariant to all possible group actions.

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. This gives constraint one forms

$$\omega_1(q) = -\sin(\theta + \phi_1)dx + \cos(\theta + \phi_1)dy - r\cos\phi_1 d\theta$$

$$\omega_2(q) = -\sin(\theta + \phi_2)dx + \cos(\theta + \phi_2)dy + r\cos\phi_2 d\theta. \quad (5.4.2)$$

These constraints are invariant under the SE(2) action given by

$$(x, y, \theta, \psi, \phi_1, \phi_2) \mapsto (x\cos\alpha - y\sin\alpha + a, x\sin\alpha + y\cos\alpha + b, \theta + \alpha, , \psi, \phi_1, \phi_2), \quad (5.4.3)$$

where (a, b, α) ∈ SE(2). The constraints determine the kinematic distribution D_q:

$$D_q = \text{span}\,\{\frac{\partial}{\partial\psi}, \frac{\partial}{\partial\phi_1}, \frac{\partial}{\partial\phi_2}, a\frac{\partial}{\partial x}, b\frac{\partial}{\partial y}, c\frac{\partial}{\partial\theta}\},$$

where a, b, and c, are given by

$$\text{a} = -r(\cos\phi_1\cos(\theta + \phi_2) + \cos\phi_2\cos(\theta + \phi_1))$$

$$\text{b} = -r(\cos\phi_1\sin(\theta + \phi_2) + \cos\phi_2\sin(\theta + \phi_1))$$

$$\text{c} = \sin(\phi_1 - \phi_2). \tag{5.4.4}$$

The tangent space to the orbits of the SE(2) action is given by

$$T_q(\text{Orb}(q)) = \text{span}\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial \theta}\}$$

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

thus given by

$$D_q \cap T_q(\text{Orb}(q)) = \text{a}\frac{\partial}{\partial x}, b\frac{\partial}{\partial y}, c\frac{\partial}{\partial \theta}. \tag{5.4.5}$$

The momentum can be constructed by choosing a section of D ∩ TOrb regarded as a bundle over Q. Since $D_q \cap T_q\text{Orb}(q)$ is one-dimensional, the section can be chosen to be

$$\xi_Q^q = \text{a}\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial \theta},$$

which is invariant under the action of SE(2) on Q. The corresponding Lie algebra element in se(2), ξ^q, is

$$\xi^q = (\text{a} + yc)e_x + (\text{b} - xc)e_y + ce\theta$$

where e_x is the basis element of the Lie algebra corresponding to translations in the x direction (and whose corresponding infinitesimal generator is $\partial/\partial x$), etc. The nonholonomic momentum map is thus given by

$$p = J^{\text{nh}}(\xi^q) = \frac{\partial L}{\partial \dot{q}^i}(\xi_Q^q)^i$$

$$= \text{ma}\dot{x} + \text{mb}\dot{y} + Jc\dot{\theta} + J_0c(\dot{\theta} + \dot{\psi}) + J_1c(\dot{\theta} + \dot{\phi}_1) + J_2c(\dot{\theta} + \dot{\phi}_2). \tag{5.4.6}$$

Here a simplification is made in which we shall also assume, namely $\phi_1 = -\phi_2$, $J_1 = J_2$. The parameters are also chosen such that $J + J_0 + J_1 + J_2 = mr^2$ (which eliminates some terms in the derivation but does not affect the essential geometry of the problem). Setting $\phi = \phi_1 = -\phi_2$, the constraints plus the momentum are

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

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

$$p = -2mr\cos^2(\phi)\cos(\theta)\dot{x} - 2mr\cos^2(\phi)\sin(\theta)\dot{y}$$

$$+ 2mr^2\cos(2\phi)\dot{\theta} + J_0\sin(2\phi)\dot{\psi}.$$

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

$$\begin{bmatrix} \cos(\theta)\dot{x} + \sin(\theta)\dot{y} \\ -\sin(\theta)\dot{x} + \cos(\theta)\dot{y} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} \frac{J_0}{2mr}\sin(2\phi)\dot{\psi} \\ 0 \\ \frac{J_0}{2mr^2}\sin^2(2\phi)\dot{\psi} \end{bmatrix} = \begin{bmatrix} \frac{-1}{2mr}p \\ 0 \\ \frac{\tan\phi}{2mr^2}p \end{bmatrix}. \qquad (5.4.7)$$

These equations have the form

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

where

$$A_{loc} = \frac{J_0}{2mr}\sin(2\phi)e_x d\psi + \frac{J_0}{2mr^2}\sin^2(\phi)e_\theta d\psi$$

$$\Gamma_{(r)} = \frac{-1}{2mr}e_x + \frac{1}{2mr^2}\tan(\phi)e_\theta.$$

These are precisely the terms which appear in the nonholonomic connection relative to the

$$\dot{p} = \frac{\partial L}{\partial \dot{q}^i}\left[\frac{d}{dt}\xi^q\right]_Q^i$$

$$= 4mr\cos(\theta)\cos(\phi)\sin(\phi)\dot{x}\dot{\phi} + 4mr\sin(\theta)\cos(\phi)\sin(\phi)\dot{y}\dot{\phi}$$

$$+ 2J_0\cos(2\phi)\dot{\phi}\dot{\psi} + 2mr^2\cos(2\phi)\dot{\theta}\dot{\phi}$$

$$-2mr\cos(\theta)cos^2(\phi)\dot{y}\dot{\theta} + 2mr\sin(\theta)cos^2(\phi)\dot{x}\dot{\theta}$$

Solving for the group velocities $\dot{x}$, $\dot{y}$, $\dot{\theta}$ from the equations which define the nonholonomic connection, the momentum equation can be rewritten as

$$\dot{p} = 2J_0 cos^2(\phi)\dot{\phi}\dot{\psi} - \tan(\phi)\, p\dot{\phi}$$

This version of the momentum equation corresponds to the coordinate form in body representation but it contains no terms which are quadratic in p, due to the fact that g^q is one dimensional.

These equations describe how paths in the base space, parameterized by r ∈ $S^1 \times S^1 \times S^1$ (in fact, the base space is $S^1 \times S^1$ if we assume $\phi_1 = -\phi_2$), are lifted to the fiber SE(2). The utility of these equations is that they greatly simplify the process of solving for the motion of the system given the base space trajectory.

Now we are ready to apply the method of reduced Lagrangian optimization to find the optimal trajectories for the snakeboard. Clearly the snakeboard is a simple nonholonomic mechanical system with symmetry as defined earlier and which also has a trivial principal bundle structure. Moreover, the control forces are only applied to the shape variables which we have full control of. Let the cost function be $C(\dot{r}) = \frac{1}{2}[(\dot{\psi})^2+(\dot{\phi})^2]$ for simplicity. We can use the method of Lagrange multipliers and Lagrangian reduction to find the necessary conditions for the optimal trajectories of the following optimal control problem:

Optimal Control Problem for the Snakeboard

Given two points $q_0, q_1 \in SE(2) \times S^1 \times S^1$, find the optimal control curves (ψ(t), ϕ (t)) ∈ $S^1 \times S^1$ that steer from q_0 to q_1 and minimize

$$\int_0^1 \frac{1}{2}[(\dot{\psi})^2 + (\dot{\phi})^2]\,dt,$$

subject to the constraints

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

$$\dot{p} = 2J_0 cos^2(\phi)\dot{\phi}\dot{\psi} - \tan(\phi)\, p\dot{\phi}$$

where

$$A_{loc} = \frac{J_0}{2mr}\sin(2\phi)e_x d\psi + \frac{J_0}{2mr^2}\sin^2(\phi)e_\theta d\psi$$

$$\Gamma_{(r)} = \frac{-1}{2mr} e_x + \frac{1}{2mr^2} \tan(\phi) e_\theta .$$

Following the general procedures in the previous section, we define a new L by

$$\mathcal{L} = \frac{1}{2}\left((\dot{\psi})^2 + (\dot{\phi})^2\right) + \lambda_a \xi^a - \frac{J_O}{2mr}\lambda_1 \sin(2\phi)\dot{\psi} + \frac{J_O}{2mr^2}\lambda_3 \sin^2(\phi)\dot{\psi}$$

$$+ \frac{J_O}{2mr}\lambda_1 p - \frac{J_O}{2mr^2}\lambda_3 \tan(\phi) p + k\dot{p} - 2J_O k \cos^2(\phi)\dot{\phi}\dot{\psi} + k \tan(\phi) p\dot{\phi}$$

where $\xi = g^{-1}\dot{g} \in$ g, $\lambda(t) \in$ g* and $k(t) \in \mathbb{R}^1$ are Lagrange multipliers. Here ξ^a and λ_a are the components of ξ and λ in the standard basis of se(2) and se(2)_ respectively.

We know that the reduced optimal curves (ψ(t), ϕ(t), $\dot{\psi}$ (t), $\dot{\phi}$(t), ξ^a(t)) must satisfy the reduced Euler Lagrangian equations for $\mathcal{L}$. After some computations, we find

$$\frac{\partial \mathcal{L}}{\partial \dot{\psi}} = \dot{\psi} - \frac{J_O}{2mr}\lambda_1 \sin(2\phi) + \frac{J_O}{2mr^2}\lambda_3 \sin^2(\phi) - 2J_O k \cos^2(\phi)\dot{\phi}$$

$$\frac{\partial \mathcal{L}}{\partial \psi} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi} - 2J_O k \cos^2(\phi)\dot{\psi} + k \tan(\phi) p$$

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{J_O}{mr}\lambda_1 \cos(2\phi)\dot{\psi} + \frac{J_O}{mr^2}\lambda_3 \sin(2\phi)\dot{\psi} - \frac{J_O}{2mr^2}\lambda_3 \sec^2(\phi) p$$

$$+ 2J_O k \cos(2\phi)\dot{\phi}\dot{\psi} + k \sec^2(\phi) p\dot{\phi}$$

$$\frac{\partial \mathcal{L}}{\partial \dot{p}} = k$$

$$\frac{\partial \mathcal{L}}{\partial p} = -\frac{J_O}{2mr}\lambda_1 + \frac{J_O}{2mr^2}\lambda_3 \tan(\phi) - k \tan(\phi)\dot{\phi}$$

$$\frac{\partial \mathcal{L}}{\partial \xi^b} = \lambda_b .$$

Substitute the above calculations into the reduced Euler Lagrangian equations and simplify, giving

$$\ddot{\psi} - \frac{J_O}{2mr}\dot{\lambda}_1 \sin(2\phi) - \frac{J_O}{mr}\lambda_1 \cos(2\phi)\dot{\phi} + \frac{J_O}{mr^2}\lambda_3 \sin(2\phi)\dot{\phi}$$

$$+ \frac{J_O}{mr^2}\dot{\lambda}_3 \sin^2(\phi) - 2J_O\dot{k} \cos^2 \phi\dot{\phi} + 2J_O k \sin(2\phi)(\dot{\phi})^2 - 2J_O k \cos^2(\phi)\ddot{\phi} = 0$$

$$\ddot{\phi} - 2J_O\dot{k} \cos^2(\phi)\dot{\psi} - 2J_O k \cos^2(\phi)\ddot{\psi} + \dot{k}\tan(\phi) p + k\tan(\phi)\dot{p}$$

$$= - \frac{J_O}{mr}\lambda_1 \cos(2\phi)\dot{\psi} + \frac{J_O}{mr^2}\lambda_3 \sin(2\phi)\dot{\psi} - \frac{J_O}{2mr^2}\lambda_3 \sec^2(\phi)\, p.$$

Also, we have

$$\dot{k} = \frac{J_O}{2mr}\lambda_1 - \frac{J_O}{2mr^2}\lambda_3 \tan(\phi) + k\tan(\phi)\dot{\phi}$$

$$\dot{\lambda}_1 = \lambda_2 \xi^3 = \lambda_2 \left(- \frac{J_O}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{1}{2mr^2}\tan(\phi) p\right)$$

$$\dot{\lambda}_2 = -\lambda_1 \xi^3 = -\lambda_1 \left(- \frac{J_O}{mr^2}\sin^2(\phi)\dot{\psi} + \frac{1}{2mr^2}\tan(\phi) p\right)$$

$$\dot{\lambda}_3 = -\lambda_2 \xi^1 = -\lambda_2 \left(\frac{J_O}{2mr}\sin(2\phi)\dot{\psi} + \frac{1}{2mr} p\right)$$

$$\dot{p} = 2J_O \cos^2(\phi)\dot{\phi}\dot{\psi} - \tan(\phi)\, p\dot{\phi}.$$

After eliminating $\dot{\lambda}_1$, $\dot{\lambda}_3$, $\dot{k}$ and $\dot{p}$ from the first set of two equations, we finally obtain

$$\ddot{\psi} - \frac{J_O}{2mr}\lambda_1 (1+3\cos(2\phi))\dot{\phi} + \frac{3J_O}{2mr^2}\lambda_3 \sin(2\phi)\dot{\phi} + J_O k \sin(2\phi)(\dot{\phi})^2 - 2J_O k \cos^2(\phi)\ddot{\phi} = 0$$

$$\ddot{\phi} - \frac{J_O}{mr}\lambda_1 \sin^2\phi\dot{\psi} + 2\frac{J_O}{mr}\lambda_1 \tan(\phi) p + \frac{1}{2mr^2}\lambda_3 p - \frac{J_O}{2mr^2}\lambda_3 \sin(2\phi)\dot{\psi} - 2J_O k \cos^2(\phi)\ddot{\psi} = 0.$$

(5.4.8)

5.5 Optimal Control on a Lie Group

Let us consider the following optimal control problem on a finite dimensional Lie group G which has been used to model various problems in several research areas (e.g. the plateball problem, and the landing tower problem). While it is possible to model this class of problems as a special case of the optimal control of nonholonomic system on a trivial principal bundle and apply reduced Lagrangian optimization, it may be useful to provide a more direct proof that uses simpler machinery.

Optimal Control Problem for a Lie Group Given a left invariant control system on G, $\dot{g} = g \cdot \xi_u$, where $\xi_u = e_0 + \sum_{i=1}^{m} u^i\ (t)\ e_i$, find the optimal controls u(·) that steer from g_0 to g_1 and $\int_0^1 L(u)$dt

Here $\{e_0, e_1, \ldots, e_m\}$ spans an (m + 1)-dimensional subspace of the whole Lie algebra $\mathfrak{g}$ of G, m+1 ≤ n = dim ($\mathfrak{g}$), u(·) is a vector valued control function with u^i (t) ∈ ℝ, L is a cost function on $\mathbb{R}^m$ which is the space of values of controls, and L(u) = $\frac{1}{2}\sum_{i=1}^{m} I_i (u^i)^2$ with $I_i > 0$.

To apply the method of Lagrangian reduction, we recast the above optimal control problem as a constrained variational problem. For simplicity of exposition, we will deal with the vector space case first where there is no e_0 term and will take up the affine case later.

Let C be the m-dimensional subspace of $\mathfrak{g}$ spanned by $\{e_0, e_1, \ldots, e_m\}$. We make the following points

(i) $\xi_u = \sum_{i=1}^{m} u^i\ (t)\ e_i$ lies in C;

(ii) if we define $L_1 = L \circ \phi$ where $L = \frac{1}{2}\sum_{i=1}^{m} I_i (u^i)^2$ with $I_i > 0$ and $\phi = (e^1, \ldots, e^m)$ with $\{e^1, \ldots, e^m\}$ as the dual basis of $\{e_1, \ldots, e_m\}$, then L_1: $C \to \mathbb{R}$ is nothing but $\frac{1}{2}$ of the square of a metric on C which is intrinsically defined and does not depend on the basis chosen;

(iii) we can extend L_1 to be half of the square of a metric $\bar{L}$ on $\mathfrak{g}$ such that $\bar{L} = L_1$ on C. As we will see, the necessary conditions for an optimal control do not depend on how the extension is done.

(iv) $\xi_u - e_0 = \sum_{i=1}^{m} u^i\ (t)\ e_i$.

Now it should be clear that the original problem is equivalent to the following constrained variational problem:

Constrained Variational Problem for Optimal Control on Lie Groups Given an m-dimensional subspace C of , find the optimal control curves ξ - $e_0 \in C$ such that g (0) = g_0, g (1) = g_1 and minimize $\int_0^1 \bar{L}(\xi - e_0)$dt.

Since we want to use the method of Langrangemultipliers to relax the constraint on the variations, we define a new Langrangian

$$\mathcal{L} = \bar{L}(\xi - e_0) + \lambda(t)(\xi - e_0) = \tilde{L}(\xi) + \tilde{\lambda}(t)(\xi) \tag{5.5.1}$$

where λ(t) lies in the annihilator C^0 of C furthermore $\tau(\xi) = \xi - e_0$, $\bar{L} = \tilde{L}\ o\ \tau$.

Theorem 5.5.1 Optimization Theorem for Nonholonomic Systems on Lie Groups. If $\bar{\xi}$ is α (regular) optimal control curve in $C + e_0 = \{\xi \in \mathfrak{g}: \xi = \xi_c + e_0, \xi_c \in C\}$, then there exists α λ(t) $\in \mathfrak{g}^*$ such that $\bar{\xi}$ satisfies the Euler-Poincare equation:

$$\frac{d}{dt}\left(\frac{\delta\tilde{L}}{\delta\xi} + \lambda\right) = \mathrm{ad}_\xi^*\left(\frac{\delta\tilde{L}}{\delta\xi} + \lambda\right) \tag{5.5.2}$$

Proof If $\bar{\xi}$(t) is an optimal control curve in $C + e_0$, then by the Lagrangian reduction method, $\bar{\xi}$(t) is a solution of the following variational problem

$$\delta\int_0^1 \mathcal{L}(\xi)\, dt = \delta\int_0^1 \tilde{L}(\xi) + \tilde{\lambda}(\xi) dt = 0$$

for some $\lambda \in \mathfrak{g}^*$, where the variations take the form $\delta\xi = \dot{\Omega} + [\xi, \Omega]$ with $= g^{-1} \cdot \delta g$ arbitrary except vanishing at the endpoints. Since

$$0 = \delta\int_0^1 \tilde{L}(\xi) + \tilde{\lambda}(\xi) dt$$

$$= \int_0^1 \left(\frac{\delta\tilde{L}}{\delta\xi}\delta\xi + \lambda(\delta\xi)\right) dt$$

$$= \int_0^1 \left(\frac{\delta\tilde{L}}{\delta\xi} + \lambda\right)(\dot{\Omega} + [\xi, \Omega])\, dt$$

$$= \int_0^1 \left(-\frac{d}{dt}\left(\frac{\delta\tilde{L}}{\delta\xi} + \lambda\right) + \mathrm{ad}_\xi^*\left(\frac{\delta\tilde{L}}{\delta\xi} + \lambda\right)\right)\Omega dt,$$

$$\mathrm{ad}_\xi^*\left(\frac{\delta\tilde{L}}{\delta\xi} + \lambda\right)$$

We conclude that $\bar{\xi}(t)$ satisfies

$$\frac{d}{dt}\left(\frac{\delta\tilde{L}}{\delta\xi}+\lambda\right)=\mathrm{ad}_{\xi}^{*}\left(\frac{\delta\tilde{L}}{\delta\xi}+\lambda\right)$$

Corollary 5.5.2 *Given a left invariant control system on G, $\dot{g} = g \cdot \xi_u$, where*

$$\xi_u = e_0 + \sum_{i=1}^{m} u^i(t)\, e_i.$$

If $\bar{\xi}(\cdot)$ is an optimal control, then

$$u^{-i}(t) = \frac{\mu_i}{I_i}$$

where i = 1, . . . ,m, and μi, i = 1, . . . ,m is the solution of the following system of differential equations

$$\dot{\mu} = C_{ji}^{k}\mu k \xi_u^j$$

where i, j, k = 0, . . . , n − 1, and where C_{ji}^{k} are the structure constants of $\mathfrak{g}$.

Proof Extend $\{e_0, e_1, \ldots, e_m\}$ to a basis $\{e_0, \ldots, e_{n-1}\}$ and let $\{e_0, \ldots, e_{n-1}\}$ be its dual basis.

(i) For i = 1, . . . ,m, and $\xi_u \in e_0 + C$, we have

$$\frac{\delta\tilde{L}}{\delta\xi_u^i} = \frac{\partial L}{u^i} = I_i u^i$$

because $\tilde{L}(\xi_u) = L \; o \; \phi \; o \; \tau\; (\xi_u) = L(u)$ and $\xi_u^i = u^i$; furthermore,

$$\lambda_i = 0, \qquad i = 1, \ldots, m$$

because λ lies in the annihilator C^o.

(ii) If we set

$$\mu_i = \frac{\delta\tilde{L}}{\delta\xi_u^i}, \; i = 1, \ldots, m,$$

and

$$\mu_i = \frac{\delta\tilde{L}}{\delta\xi_u^i} + \lambda_i, \; i = m+1, \ldots, n-1 \; 0,$$

and write out the Euler-Poincare equation using the above coordinates, we will get the desired system of differential equations. [11]

Remarks

1. From the above computations, we can see that the necessary conditions for an optimal control ū (·) depend only on L and have nothing to do with how the extension is done, because not only
   $u^i(t) = \mu_i(t)/I_i$, but also $\dot{\mu}_i = C_{ji}^{k} \mu k \xi_u^j$ do not depend on $\bar{L}$.
2. The necessary conditions given in the above Corollary are the same as those in Krishnaprasad [1994]:

$$u^i = \frac{\mu_i}{I_i} \qquad\qquad i = 1, \dots, m,$$

$$\dot{\mu}_i = - \mu_k C_{ij}^k \frac{\delta h}{\delta \mu_i} \qquad i, j, k = 0, \dots, n-1,$$

where

$$h = \mu_0 + \frac{1}{2} \sum_{i=1}^{m} \frac{{\mu_i}^2}{I_i}.$$

This is because $C_{ji}^k = - C_{ij}^k$ and

$$\frac{\delta h}{\delta \mu_i} = \left.\begin{cases} 1 & j = 0 \\ \frac{\mu_j}{I_j} = u^j & j = 1, \dots, m \\ 0 & j = m+1, \dots, n-1 \end{cases}\right\} = \xi_u^j$$