Time-optimal control of two simple mechanical systems with three degrees of freedom and two inputs*

A. Theo Coombs†

Contents

---

*Report for project in fulfilment of requirements for MSc

†Graduate student, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
Email: `theo@mast.queensu.ca`

Research partially supported by a fellowship from Ontario Graduate Scholarships in Science and Technology.

1. Introduction

To take advantage of the applications to mechanical systems that mathematics has to offer, one would wish to have a satisfactory understanding of the theory involved. In this paper we try to build exactly that, and tackle the problem of time-optimal control of two simple mechanical systems: the robotic leg and the planar rigid body. Both systems have uncontrolled motions that are described by a kinetic energy Lagrangian function which is determined by the associated Riemannian metric. Each of these metrics offers us its Levi-Civita affine connection to work with. We also mention that both systems are shown to be controllable in [Lewis and Murray 1999], so the question of existence of solutions will not be of concern.

Our goal will be to understand the advances that lead up to a maximum principle for affine connection control systems given by Bullo and Lewis [2005, Chapter S4], and which we specifically use for time-optimization. To start, we look briefly at some elementary results from the calculus of variations and definitions from affine differential geometry, and establish a link between the two. Next we take the path offered by Sussmann and Willems [1997] that leads us from some well-known necessary conditions for minima in the calculus of variations, to the celebrated maximum principle of optimal control theory. Then the flavor becomes more geometrical in nature, as we examine affine connection control systems and the splitting of fibres in higher-order tangent and cotangent bundles. This gives us enough insight to state two versions of the maximum principle, one being a result for the more general control affine systems that follows from the work of Sussmann [1998], and the other a special case for affine connection control systems. From the latter, another result is found specifically for systems in which the norm of the controls is bounded.

After studying all of the newly encountered concepts, we apply them to our two mechanical systems. We find the equations of motion for each system and compute the Hamiltonian equations supplied by the maximum principle. Then we apply the maximum principle for affine connection control systems and find our controls in terms of a one-form field along time-optimal solutions. To finish we examine a special "singular" case for the planar rigid body, when the maximum principle does not allow us to determine the optimal controls. For this system, the time-optimal singular extremals can be completely described.

2. The calculus of variations and affine connections

One of the earliest known optimal control problems, according to Sussmann and Willems [1997], dates to about 1696 when Johann Bernoulli posed the **brachystochrone problem** which reads:

> *If in a vertical plane two points A and B are given, then it is required to specify the orbit AMB of the movable point M, along which it, starting from A, and under the influence of its own weight, arrives at B in the shortest possible time. So that those who are keen of such matters will be tempted to solve this problem, it is good to know that it is not, as it may seem, purely speculative and without practical use. Rather it even appears, and this may be hard to believe, that it is very useful also for other branches of science than mechanics. In order to avoid a hasty conclusion, it should be remarked that the straight line*

is certainly the line of shortest distance between A and B, but it is not the one which is traveled in the shortest time. However, the curve AMB – which I shall divulge if by the end of this year nobody else has found it – is very well known among geometers.

Within a year solutions were submitted by Johann Bernoulli and others such as Newton, Leibniz, Tschirnhaus, l'Hôpital, and Johann's brother, Jakob Bernoulli. The brachystochrone problem starts us exploring a road that leads through the necessary conditions for extremizing functionals, up to the widely applicable maximum principle. The solutions to certain optimization problems turn out to be geodesics, so there must be a bridge that can be taken to the area of affine differential geometry. This is exhibited in Section 2.2.

2.1. Necessary conditions for optimization in the calculus of variations. The calculus of variations is a classical subject in the area of applied mathematics. The following is one of the interesting problems with which it deals.

2.1 Problem: Find a function $x_0(t)$ that minimizes the functional

$$J(x) = \int_a^b L(t, x(t), \dot{x}(t))\, \mathrm{d}t,$$

over all curves $x : [a, b] \to \mathbb{R}^n$ such that $x(t) \in C^\infty[a, b]$ and $L(t, x, v) \in C^\infty([a, b]\times\mathbb{R}^n\times\mathbb{R}^n)$, where $x(a) = x_0$ and $x(b) = x_1$.

We ask that the Lagrangian or cost function $L(t, x, v)$ and the class of available functions from which our minimizer can be chosen be those which are of type C^∞. Weaker hypotheses are possible, however for what is to come, this assumption will suffice and offer simplicity. This problem is at the heart of classical optimization theory, and thus turns out to be a good place for our topic of interest to begin.

We now mention three necessary conditions that must be satisfied for an extremal to be a minimum value of the functional $J(x)$. Two of them will be referred to later when developing a stronger statement that we will use to investigate our two mechanical systems. More general statements can be found in [Ewing 1985], when the class of available curves is extended to include those with discontinuous derivatives.

Euler's necessary condition. One of the main results in the calculus of variations is a necessary condition formulated by Euler around the year 1744. It comes from a simple idea in elementary calculus, that if an extremal exists, the derivative of the function being maximized or minimized must vanish. We will be adopting the ***summation convention*** where summation is implied over two identical indices occurring in the same term, one a subscript and one a superscript. The result can be summed up in the following theorem.

2.2 Theorem: (Euler's Necessary Condition) *If $x_0(t)$ solves Problem 2.1 then $x_0(t)$ must satisfy the Euler-Lagrange equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v}(t, x_0(t), \dot{x}_0(t)) = \frac{\partial L}{\partial x}(t, x_0(t), \dot{x}_0(t)), \tag{2.1}$$

for all $t \in [a, b]$.

Proof: We consider arbitrary variations of the minimizing function $x_0(t) + \lambda h(t)$, where $h(t) \in C^\infty[a,b]$ and $h(a) = h(b) = 0$, and take $g(\lambda) = J(x_0(t) + \lambda h(t))$. Since $L(t,x,v)$ is differentiable and we take the interval $[a,b]$ to be compact, then $g(\lambda)$ is differentiable. If $x_0(t)$ solves the minimization problem, then $g(0)$ must be a minimum value of $g(\lambda)$ and therefore it is necessary that $g'(0) = 0$. Differentiating g with respect to λ and evaluating at $\lambda = 0$,

$$\begin{aligned} g'(0) &= \int_a^b \frac{\partial}{\partial \lambda}\Big|_{\lambda=0} L(t, x_0(t) + \lambda h(t), \dot{x}_0(t) + \lambda \dot{h}(t))\, \mathrm{d}t \\ &= \int_a^b \left(\frac{\partial L}{\partial x^i}(t, x_0(t), \dot{x}_0(t)) h^i(t) + \frac{\partial L}{\partial v^i}(t, x_0(t), \dot{x}_0(t)) \dot{h}^i(t) \right) \mathrm{d}t \\ &= \int_a^b \left(\frac{\partial L}{\partial x^i}(t, x_0(t), \dot{x}_0(t)) h^i(t) - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial v^i}(t, x_0(t), \dot{x}_0(t)) h^i(t) \right) \mathrm{d}t \\ &\quad + \frac{\partial L}{\partial v^i}(t, x_0(t), \dot{x}_0(t)) h^i(t) \Big|_{t=a}^{t=b}, \end{aligned}$$

where the last step includes integration by parts on the second term. Using the endpoint conditions, $h^i(a) = h^i(b) = 0$ for $i = 1, \ldots, n$, we have

$$g'(0) = \int_a^b \left(\frac{\partial L}{\partial x^i}(t, x_0(t), \dot{x}_0(t)) - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial v^i}(t, x_0(t), \dot{x}_0(t)) \right) h^i(t)\, \mathrm{d}t.$$

And since $g'(0) = 0$ and $h(t)$ was chosen arbitrarily throughout $[a,b]$, subject to the endpoint constraints, we obtain equation (2.1), as required.□

The Euler-Lagrange equation gives us a necessary condition for an extremal to occur, so once we obtain such candidates, we would want to check if they are a maximum or a minimum. We will be mainly dealing with minima throughout our discussions.

The Wieirstrass excess function. The next necessary condition that we should state involves the ***excess function*** defined by

$$E(t, x, r, q) = L(t, x, q) - L(t, x, r) - (q - r) \frac{\partial L}{\partial v}(t, x, r). \tag{2.2}$$

This function measures the difference between $L(t,x,q)$ and its first order Taylor polynomial based at the point (t,x,r).

2.3 Theorem: (Wieirstrass) *If $x_0(t)$ solves Problem 2.1 then*

$$E(t, x_0(t), \dot{x}_0(t), q) \geq 0,$$

for every $t \in [a,b]$ and for every $q \in \mathbb{R}$.

We omit a proof of this statement, but refer our readers to [Ewing 1985], or any suitable text on variational calculus.

Legendre's necessary condition. Another very important result comes from Legendre around 1786. It concerns a necessary condition for a minimum to occur, and similar to Euler's necessary condition, can be thought of in terms of ideas from elementary calculus. This time however, it is the fact that the second derivative evaluated at the point admitting a minimum must be non-negative.

2.4 Theorem: (Legendre's necessary condition) *If $x_0(t)$ solves Problem 2.1 then $x_0(t)$ must satisfy*

$$\frac{\partial^2 L}{\partial v^2}(t, x_0(t), \dot{x}_0(t)) \geq 0, \tag{2.3}$$

for all $t \in [a, b]$.

Proof: We apply Taylor's Formula with a second-order remainder to $L(t, x, q)$ about the point r,

$$L(t,x,q) = L(t,x,r) + (q-r)\frac{\partial L}{\partial v}(t,x,r) + \frac{(q-r)^2}{2}\frac{\partial^2 L}{\partial v^2}(t,x,\theta(q-r)),$$

for some $\theta \in (0,1)$. Using the excess function, equation (2.2), we find that

$$E(t,x,r,q) = \frac{(q-r)^2}{2}\frac{\partial^2 L}{\partial v^2}(t,x,\theta(q-r)).$$

And from Theorem 2.3, when we take $r = \dot{x}(t)$ and $q = \frac{1+\theta}{\theta}\dot{x}(t)$, the statement of the current theorem follows.□

2.2. Affine connections in the theory of geodesics.

Differential geometry provides the applied mathematician with a lot of new concepts and techniques, some of which offer a significant number of applications to variational problems. When looking for solutions to optimization problems, it is nice to know that they can sometimes turn out to be simply geodesics on a manifold supplied with an affine connection. Therefore we build a direct relationship between certain variational problems and geodesic theory.

Definitions. We first present some affine differential geometry that will be used throughout our discussion. The motivation here is simply for notational purposes. Most of the definitions come from [Kobayashi and Nomizu 1963a] and [Kobayashi and Nomizu 1963b], except for the adjoint forms of torsion and curvature, and the adjoint Jacobi equation, which may be found in [Bullo and Lewis 2005, Chapter S4].

2.5 Definition: An ***affine connection*** on a manifold Q is an assignment to each pair of vector fields X and Y on Q, a vector field $\nabla_X Y$, and the assignment should satisfy the properties:

1. the map $(X, Y) \mapsto \nabla_X Y$ is $\mathbb{R}$-bilinear,
2. $\nabla_{fX} Y = f\nabla_X Y$ for $f \in C^\infty(Q)$, and
3. $\nabla_X fY = f\nabla_X Y + (\mathscr{L}_X f)Y$ for $f \in C^\infty(Q)$,

where $\mathscr{L}_X f$ is the Lie derivative of the function f with respect to the vector field X.

The vector field $\nabla_X Y$ is called the ***covariant derivative*** of Y with respect to X. If $(q^1, \ldots, q^n)$ are coordinates for Q, then $\nabla_X Y$ in coordinates is given by

$$\nabla_X Y = \left(\frac{\partial Y^i}{\partial q^j}X^j + \Gamma^i_{jk}X^jY^k\right)\frac{\partial}{\partial q^i}, \tag{2.4}$$

where the n^3 functions Γ^i_{jk} are called the ***Christoffel symbols*** for the affine connection ∇ defined by

$$\nabla_{\frac{\partial}{\partial q^j}} \frac{\partial}{\partial q^k} = \Gamma^i_{jk} \frac{\partial}{\partial q^i}.$$

This tells us how to covariantly differentiate a vector field. We define the covariant derivative of a function to be its Lie derivative

$$\nabla_X f = \mathscr{L}_X f.$$

To find the covariant derivative of a one-form Λ, we first apply the one-form to an arbitrary vector field Y. We would like $\nabla_X \Lambda$ to have the property

$$\nabla_X(\Lambda(Y)) = (\nabla_X \Lambda)(Y) + \Lambda(\nabla_X Y),$$

which is basically the product rule from elementary calculus. Indeed, this is how we define the covariant derivative of a one-form applied to a vector field,

$$(\nabla_X \Lambda)(Y) = \mathscr{L}_X(\Lambda(Y)) - \Lambda(\nabla_X Y).$$

To obtain the coordinate expression, we use equation (2.4) and the fact that the Lie derivative in coordinates is $\mathscr{L}_X f = \frac{\partial f}{\partial q^i} X^i$. One may then verify that

$$\nabla_X \Lambda = \left(\frac{\partial \Lambda_i}{\partial q^j} X^j - \Gamma^j_{ki} X^k \Lambda_j \right) \mathrm{d}q^i. \tag{2.5}$$

It is also possible to covariantly differentiate an arbitrary (r, s)-tensor. To do this, we take such a tensor $t \in T^r_s(TM)$ and thus its covariant derivative with respect to a vector field X will be of the same type: $\nabla_X t \in T^r_s(TM)$. A (r, s)-tensor takes as its arguments, r one-forms and s vector fields. If $\{\Lambda^1, \ldots, \Lambda^r\}$ is a set of one-forms and $\{Y_1, \ldots, Y_s\}$ is a set of vector fields, then we wish to have

$$\begin{aligned}\nabla_X(t(\Lambda^1, \ldots, \Lambda^r, Y_1, \ldots, Y_s)) = (\nabla_X t)(\Lambda^1, \ldots, \Lambda^r, Y_1, \ldots, Y_s) \\ - \sum_{i=1}^{r} t(\Lambda^1, \ldots, \nabla_X \Lambda^i, \ldots, \Lambda^r, Y_1, \ldots, Y_s) \\ - \sum_{j=1}^{s} t(\Lambda^1, \ldots, \Lambda^r, Y_1, \ldots, \nabla_X Y_j, \ldots, Y_s).\end{aligned}$$

It is then possible for one to solve for $\nabla_X t$ and obtain the covariant derivative of an arbitrary (r, s)-tensor.

A useful property of affine connections is that they allow us to differentiate vector fields and one-forms along curves. Consider a curve $c(t)$ on Q, a vector field $v(t)$ and one-form field $\lambda(t)$ along c. Let X and Y be vector fields such that X has c as an integral curve, i.e., $c'(t) = X(c(t))$, and $v(t) = Y(c(t))$. Then we define the ***covariant derivative of v along c*** to be the vector field along c given by

$$\nabla_{c'(t)} v(t) = \nabla_X Y(c(t)).$$

In coordinates,

$$\nabla_{c'(t)}v(t) = \big(\dot{v}^i(t) + \Gamma^i_{jk}\dot{q}^j v^k(t)\big)\frac{\partial}{\partial q^i}, \tag{2.6}$$

which comes from equation (2.4).

Now let X be a vector field and Λ a one-form such that X has c as an integral curve, i.e., $c'(t) = X(c(t))$, and $\lambda(t) = \Lambda(c(t))$. Then we define the ***covariant derivative of $\boldsymbol{\lambda}$ along $\boldsymbol{c}$*** by

$$\nabla_{c'(t)}\lambda(t) = \nabla_X\Lambda(c(t)).$$

In coordinates

$$\nabla_{c'(t)}\lambda(t) = \big(\dot{\lambda}_i(t) - \Gamma^j_{ki}\dot{q}^k\lambda_j(t)\big)\mathrm{d}q^i, \tag{2.7}$$

which can be verified from equation (2.5).

On a manifold supplied with an affine connection, a certain class of curves is distinguished. These are called ***geodesics*** and we will see them again when trying to establish links with variational problems and in the drift motion of state trajectories for mechanical systems. A ***geodesic*** $c(t)$ of the affine connection satisfies

$$\nabla_{c'(t)}c'(t) = 0.$$

One verifies that a geodesic in coordinates satisfies

$$\ddot{q}^{\,i} + \Gamma^i_{jk}\dot{q}^j\dot{q}^k = 0. \tag{2.8}$$

Since the geodesic equation is a second-order differential equation, it defines a second-order vector field Z on TQ called the ***geodesic spray*** of the affine connection ∇, given in coordinates by

$$Z = v^i\frac{\partial}{\partial q^i} - \Gamma^i_{jk}v^j v^k\frac{\partial}{\partial v^i}, \tag{2.9}$$

so that the coefficients of $\frac{\partial}{\partial q^i}$ and $\frac{\partial}{\partial v^i}$ are $\dot{q}^i$ and $\ddot{q}^i$, respectively. This vector field will play a significant role in what is to come.

We now look at four tensor fields that give us an idea of the "shape" of our manifold equipped with its affine connection. The second definition may be somewhat familiar from elementary vector calculus.

2.6 Definition: The ***torsion*** T for ∇ is a $(1,2)$-tensor field given by

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y],$$

so, in coordinates

$$T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj}.$$

One may verify that the torsion tensor is bilinear in X and Y. For $u_q \in T_qQ$ and $\alpha_q \in T^*_qQ$, we define the ***adjoint torsion*** $T^*(\alpha_q, u_q) \in T^*_qQ$ by

$$\langle T^*(\alpha_q,u_q); w_q\rangle = \langle \alpha_q; T(w_q,u_q)\rangle, \quad w_q \in T_qQ.$$

2.7 Definition: The ***curvature*** R for ∇ is a $(1,3)$-tensor field given by

$$R(X,Y)W = \nabla_X \nabla_Y W - \nabla_Y \nabla_X W - \nabla_{[X,Y]} W,$$

so that in coordinates

$$R^i_{jkl} = \frac{\partial \Gamma^i_{lj}}{\partial q^k} - \frac{\partial \Gamma^i_{kj}}{\partial q^l} + \Gamma^i_{km}\Gamma^m_{lj} - \Gamma^i_{lm}\Gamma^m_{kj}.$$

Of course, the curvature tensor is multilinear in X, Y, and W. For $u_q, v_q \in T_qQ$ and $\alpha_q \in T^*_qQ$, we define the ***adjoint curvature*** $R^*(\alpha_q, u_q)v_q \in T^*_qQ$ by

$$\langle R^*(\alpha_q, u_q)v_q; w_q\rangle = \langle \alpha_q; R(w_q, u_q)v_q\rangle, \quad w_q \in T_qQ.$$

The affine connections that we will be working with when investigating the two mechanical systems are ***Levi-Civita*** connections, so we will need to know their properties. A ***Riemannian metric*** is a symmetric positive-definite $(0,2)$-tensor that gives us a notion of the distance between points on our manifold, and the unique Levi-Civita connection supplied with such a metric g is denoted by $\overset{g}{\nabla}$.

2.8 Definition: Let g be a Riemannian metric on a manifold Q. The ***Levi-Civita connection*** is the unique affine connection satisfying the following properties:

1. $\overset{g}{\nabla}_X g = 0$ for all vector fields $X \in Q$,
2. the torsion $T(X,Y) = 0$ for all vector fields $X, Y \in Q$.

The Christoffel symbols for the Levi-Civita connection may be shown to be

$$\Gamma^i_{jk} = \frac{1}{2} g^{il}\left(\frac{\partial g_{lj}}{\partial q^k} + \frac{\partial g_{lk}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^l}\right).$$

The next two definitions characterize a way to measure variations of geodesics. These will be necessary to come to a version of the maximum principle for affine connection control systems in Section 3.3.

2.9 Definition: A vector field $\xi(t)$ along a geodesic $c(t)$ is a ***Jacobi field*** if it satisfies the ***Jacobi equation***

$$\nabla^2_{c'(t)}\xi(t) + R(\xi(t), c'(t))c'(t) + \nabla_{c'(t)}(T(\xi(t), c'(t))) = 0. \tag{2.10}$$

2.10 Definition: A one-form field $\alpha(t)$ along a geodesic $c(t)$ is an ***adjoint Jacobi field*** if it satisfies the ***adjoint Jacobi equation***

$$\nabla^2_{c'(t)}\alpha(t) + R^*(\alpha(t), c'(t))c'(t) - T^*(\nabla_{c'(t)}\alpha(t), c'(t)) = 0. \tag{2.11}$$

Relations to variational problems. We now try to establish some relationships between variational problems and affine connections. To start we will present an example about the two-dimensional sphere $\mathbb{S}^2$. The geodesics turn out to be the ***great circles***, and this will be demonstrated in two ways.

2.11 Example: To find the geodesics, we must minimize the distance between two points on the sphere $\mathbb{S}^2$. It can be shown that equivalently, we may minimize the square of the distance,

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2,$$

where $(x, y, z) \in \mathbb{R}^3$ are restricted to $\mathbb{S}^2$. Now the equation of a sphere can be written as

$$\begin{aligned} x = \sin\theta\cos\phi, \quad y = \sin\theta\sin\phi, \quad z = \cos\theta, \\ \Longrightarrow \mathrm{d}s^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2. \end{aligned}$$

So we take our Lagrangian in coordinates θ and ϕ,

$$L((\theta, \phi), (\dot{\theta}, \dot{\phi})) = \dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2.$$

Now the terms of equation (2.1), the Euler-Lagrange equation, are

$$\begin{aligned} \frac{\partial L}{\partial x} &= \begin{bmatrix} 2\sin\theta\cos\theta\,\dot{\phi}^2 \\ 0 \end{bmatrix}, \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v} &= \begin{bmatrix} 2\ddot{\theta} \\ 4\sin\theta\cos\theta\,\dot{\theta}\,\dot{\phi} + 2\sin^2\theta\,\ddot{\phi} \end{bmatrix}. \end{aligned}$$

On equating the components of each term, we find the equations of the geodesics on the sphere,

$$\begin{aligned} \ddot{\theta} - \sin\theta\cos\theta\,\dot{\phi}^2 = 0, \\ \ddot{\phi} + 2\cot\theta\,\dot{\theta}\,\dot{\phi} = 0. \end{aligned} \tag{2.12}$$

Let us find these geodesics from equation (2.8). Using the same equations for the sphere from before, we compute the metric to be

$$\begin{aligned} g_{\mathbb{S}^2} &= (\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z)\,|\,\mathbb{S}^2 \\ &= \mathrm{d}\theta \otimes \mathrm{d}\theta + \sin^2\theta\,\mathrm{d}\phi \otimes \mathrm{d}\phi. \end{aligned}$$

Therefore $g_{\theta\theta} = 1$, $g_{\phi\phi} = \sin^2\theta$, and $g_{\theta\phi} = g_{\phi\theta} = 0$, and since we are dealing with a Levi-Civita connection, we find the non-zero Christoffel symbols from Definition 2.8:

$$\Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta \quad \text{and} \quad \Gamma^{\phi}_{\phi\theta} = \Gamma^{\phi}_{\theta\phi} = \cot\theta.$$

Using the coordinate expression for a geodesic, equation (2.8), we find the same equations of motion as those given before in equations (2.12).

The above example shows how geodesics can turn out to be the solutions for certain variational problems. Now let us look at the generalization.

2.12 Proposition: *The Euler-Lagrange equation for the Lagrangian*

$$L(v_q) = \frac{1}{2} g(v_q, v_q)$$

is equivalent to the geodesic equation

$$\overset{g}{\nabla}_{c'(t)} c'(t) = 0,$$

for a Levi-Civita connection g.

Proof: We choose some arbitrary coordinates for $c(t)$ to be $q^i(t)$ where $i = 1, \dots, n$. Differentiating the Lagrangian $L = \frac{1}{2} g_{jk} \dot{q}^j \dot{q}^k$ first with respect to q^l and then with respect to $\dot{q}^l$ and t, and remembering that g is symmetric in its indices,

$$\begin{aligned}
\frac{\partial L}{\partial q^l} &= \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial q^l}\dot{q}^j\dot{q}^k\right), \\
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^l} &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(g_{lk}\dot{q}^k + g_{jl}\dot{q}^j\right) \\
&= \frac{1}{2}\left(\frac{\partial g_{lj}}{\partial q^k} + \frac{\partial g_{lk}}{\partial q^j}\right)\dot{q}^j\dot{q}^k + g_{il}\ddot{q}^i.
\end{aligned}$$

Now using the Euler-Lagrange equation and multiplying through by the inverse of the metric g^{il}, we arrive at the equation

$$\ddot{q}^i + \frac{1}{2} g^{il}\left(\frac{\partial g_{lj}}{\partial q^k} + \frac{\partial g_{lk}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^l}\right)\dot{q}^j\dot{q}^k = 0,$$

which is simply the coordinate expression for $\overset{g}{\nabla}_{c'(t)} c'(t) = 0$ with the Christoffel symbols Γ^i_{jk} for the Levi-Civita connection. From this we see that the reverse implication is also straightforward.□

Thus we see precisely how geodesics for Levi-Civita connections have a direct link to variational theory.

3. The maximum principle

When we want to find a function that gives a minimum value to the functional $J(x)$, we can try to use the theorems we stated from the calculus of variations. However, there is a statement that encompasses the three theorems given in Section 2.1, and holds even more information. This is the maximum principle from [Pontryagin, Boltyanskiĭ, Gamkrelidze, and Mishchenko 1986] and we will see how the development of this theorem involves the necessary conditions of the calculus of variations as shown by Sussmann and Willems [1997], and then look at another version specifically for affine connection control systems. From this, we find a statement that enables us to find the controls needed for time-optimization.

3.1. From Euler-Lagrange to the maximum principle.

For the problems that we will be dealing with, the Lagrangian will not be an explicit function of time. Thus the problem that we will now be considering is the following:

3.1 Problem: Find a function $x_0(t)$ that minimizes the functional

$$J(x) = \int_a^b L(x(t), \dot{x}(t))\,\mathrm{d}t,$$

over all curves $x : [a, b] \to \mathbb{R}^n$ such that $x(t) \in C^\infty[a, b]$ and $L(x, v) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$, where $x(a) = x_0$ and $x(b) = x_1$.

Two necessary conditions for $(x(t), \dot{x}(t))$ to minimize J are

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v}(x(t), \dot{x}(t)) = \frac{\partial L}{\partial x}(x(t), \dot{x}(t)) \quad \text{and} \quad \frac{\partial^2 L}{\partial v^2}(x(t), \dot{x}(t)) \geq 0, \tag{3.1}$$

which come from Theorems 2.2 and 2.4. Given the Lagrangian $L(x, v)$, we define the ***Hamiltonian***

$$H(x, p, v) = p \cdot v - L(x, v). \tag{3.2}$$

If $x(t)$ is a solution to equations (3.1), and if $p(t)$ is defined by $p(t) = \frac{\partial L}{\partial v}(x(t), \dot{x}(t))$, then by direct calculation we have

$$\begin{aligned}
&\frac{\partial H}{\partial x}(x(t), p(t), \dot{x}(t)) = -\dot{p}(t), \\
&\frac{\partial H}{\partial p}(x(t), p(t), \dot{x}(t)) = \dot{x}(t), \\
&\frac{\partial H}{\partial v}(x(t), p(t), \dot{x}(t)) = 0.
\end{aligned}$$

Since H is equal to $-L$ plus a linear function of v, then the Legendre condition tells us that $\frac{\partial^2 H}{\partial v^2}(x(t), p(t), \dot{x}(t)) \leq 0$, so H must be a maximum at $(x(t), \dot{x}(t))$. This leads us to the following theorem:

3.2 Theorem: *If $(x(t), \dot{x}(t))$ is a solution to Problem 3.1 then there exists $p(t)$, a one-form field along $x(t)$ such that*

1. $\dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t), \dot{x}(t))$,
2. $\dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t), \dot{x}(t))$,
3. $H(x(t), p(t), \dot{x}(t)) = \max\limits_{v \in T_x M} H(x(t), p(t), v)$,

where $H(x, p, v) = p \cdot v - L(x, v)$, $x \in M$ and $v \in T_x M$.

Thus we see how the maximum of the Hamiltonian is suggested by the Euler-Lagrange equations. We are interested in how such a result can be modified to include systems whose admissible curves are restricted by a parameterized differential equation. Thus we consider how to apply the Hamiltonian equations to control systems.

Let $U \subset \mathbb{R}^m$ be a fixed subset. If $u : [a, b] \to U$ is a bounded and measurable function, and $x(t)$ is a curve on the state manifold M, then in a control system the state trajectory $x(t)$ changes according to the differential equation

$$\dot{x}(t) = f(x(t), u(t)). \tag{3.3}$$

We can change $u(t)$ to try to make the system behave the way we want. An example of a nonlinear control system is

$$\dot{x}(t) = f_0(x(t)) + u^a(t) f_a(x(t)),$$

where $f_0(x)$ is the drift term or uncontrolled part and $f_a(x)$ determines how the controls u^a act on the system. We shall study this system later.

Equation (3.3) puts a restriction on our admissible class of curves. Essentially, $u(t)$ parameterizes the set of available velocities. A convenient, but naive, thing to do is to replace v in our Hamiltonian with $f(x, u)$. This also requires us to redefine our Lagrangian, so that it is a function of $x(t)$ and $u(t)$. Thus our new ***control*** Hamiltonian becomes

$$H(x, p, u) = p \cdot f(x, u) - p_0 L(x, u), \tag{3.4}$$

where $L(x, u)$ is the new Lagrangian (more than likely different from $L(x, v)$). The constant $p_0 \in \{0, 1\}$ is known as the ***abnormal multiplier*** and basically switches the dependence of the Hamiltonian on the Lagrangian, on or off. The necessity of the introduction of this constant is not apparent, but it must be done to ensure accurate statements.

To apply the Hamiltonian equations in this setting, we must first state the optimal control problem.

3.3 Problem: Find a pair $(u(t), x(t))$ that minimizes the functional

$$J(u, x) = \int_a^b L(x(t), u(t))\, \mathrm{d}t,$$

subject to $\dot{x}(t) = f(x(t), u(t))$, $x(a) = x_0$ and $x(b) = x_1$, where $x(t)$ is a curve on the state manifold M, $u : [a, b] \to U \subset \mathbb{R}^m$ is a measurable function, and $L(x, u) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^m)$.

Our curves will arise from the control $u(t)$, since we are choosing the control which gives a family of curves from the equation $\dot{x}(t) = f(x(t), u(t))$. Now we may use the results of the previous theorem to state a necessary condition found in [Pontryagin, Boltyanskiĭ, Gamkrelidze, and Mishchenko 1986] to minimize J with the restriction of equation (3.3), the control equation.

3.4 Theorem: (Maximum Principle) *A necessary condition for the pair $(u(t), x(t))$ to solve Problem 3.3 is that there exist a one-form field $p(t)$ along $x(t)$ and a constant $p_0 \in \{0, 1\}$ such that*

1. $(p(t), p_0) \neq (0, 0)$ *for all* $t \in [a, b]$,
2. $\dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t), u(t))$ *and* $\dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t), u(t))$ *for all* $t \in [a, b]$,
3. $H(x(t), p(t), u(t)) = \max_{\tilde{u} \in U} H(x(t), p(t), \tilde{u})$ *for all* $t \in [a, b]$,
4. $H(x, p, u) = p \cdot f(x, u) - p_0 L(x, u)$ *is constant almost everywhere along solutions and if we allow the endpoints to vary, this constant can be chosen to be zero.*

3.2. The geometry of the drift vector field. Let us examine the geometry that arises from the Hamiltonian equations. We will be working toward an understanding of how integral curves of the tangent and cotangent lifts of the geodesic spray are related to the Jacobi and adjoint Jacobi fields.

Tangent and cotangent lifts of general vector fields. We now consider the nonlinear control system characterized by the differential equation

$$\dot{x}(t) = f_0(x(t)) + u^a(t) f_a(x(t)), \tag{3.5}$$

where $x(t)$ is a curve in the state space M and $u : I \to U \subset \mathbb{R}^m$. To minimize the functional $J(u,x) = \int_a^b L(x,u)\,\mathrm{d}t$ with the restriction of equation (3.5), the control equation, we form the control Hamiltonian

$$H(x,p,u) = p(f_0(x) + u^a f_a(x)) - p_0 L(x,u).$$

Now, according to the maximum principle, a necessary condition for the pair $(u(t), x(t))$ to solve the minimization problem is that there exist a one-form $p(t)$ along $x(t)$ such that $(x(t), p(t))$ satisfies

$$\dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t), u(t)) \quad \text{and} \quad \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t), u(t)).$$

These equations imply that on the cotangent bundle T^*M we have

$$\dot{x}^i = f_0^i + u^a f_a^i \quad \text{and} \quad \dot{p}_i = -\frac{\partial f_0^j}{\partial x^i} p_j - \frac{\partial f_a^j}{\partial x^i} p_j u^a + p_0 \frac{\partial L}{\partial x^i}(x,u).$$

The first equation simply implies that $x(t)$ satisfies the control equation. The second equation for $p(t)$ is not coordinate invariant, and therefore has no meaning by itself, but only when it is considered with the first. It is the adjoint equation, but without enough structure, we can't really consider it alone. From these two equations, we are mainly interested in terms involving the uncontrolled part since these contain the affine differential geometry. This leads us to study the ***cotangent lift*** of the vector field f_0 on T^*M defined by

$$f_0^{T^*} = f_0^i \frac{\partial}{\partial x^i} - \frac{\partial f_0^j}{\partial x^i} p_j \frac{\partial}{\partial p_i}.$$

It turns out to be easier for us to investigate a related object on the tangent bundle TM. Therefore we first look at the ***tangent lift*** of the vector field f_0 defined by

$$f_0^T = f_0^i \frac{\partial}{\partial x^i} + \frac{\partial f_0^i}{\partial x^j} v^j \frac{\partial}{\partial v^i},$$

so that we may generate some useful intuition to help us later.

We will see that when we take f_0 to be the geodesic spray of our affine connection, there exists a relation between f_0^T and solutions of the Jacobi equation. This motivates us to find a similar relation for $f_0^{T^*}$. So our current purpose will be to understand the geometry of these vector fields in the general situation.

We will show that f_0^T measures variations of perturbations of the initial condition, thus showing how nearby solutions vary. We now ask for an integral curve of f_0^T, and the answer lies in the following theorem.

3.5 Theorem: *Let $c(t)$ be an integral curve for f_0, that is a solution to $c'(t) = f_0(c(t))$, and let $c_s(t)$ be a one-parameter family of integral curves for $s \in (-\varepsilon, \varepsilon)$ with $c_0(t) = c(t)$. Define $v_0 = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} c_s(0) \in T_{c(0)}M$. So v_0 gives the change in how the initial condition $c(0)$ varies. Now define*

$$v(t) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} c_s(t).$$

Then $v(t)$ is the integral curve for f_0^T with initial condition v_0.

Proof: First define the flow along $c(t)$ as $F(x,t)$ so that $F(x_0,t) = c(t)$ where $x_0 = c(0)$ and let $x_0^s = c_s(0)$. Now $F(x_0^s,t) = c_s(t)$, and so

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} c_s(t) = \frac{\partial F}{\partial x^i}(x_0,t)v_0^i,$$

where $v_0^i = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}(x_0^s)^i$. Also, $c_s'(t) = \frac{\mathrm{d}}{\mathrm{d}t}F(x_0^s,t) = f_0(c_s(t))$. So $v(t)$ is a curve in TM such that

$$v : t \mapsto \left(F(x_0,t), \frac{\partial F}{\partial x^i}(x_0,t)v_0^i \right),$$

and since $\frac{\mathrm{d}}{\mathrm{d}t}F(x_0,t) = f_0(c(t))$ and $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial x^i}(x_0,t) = \frac{\partial f_0}{\partial x^i}(c(t))$, we have

$$\frac{\mathrm{d}v}{\mathrm{d}t} : t \mapsto \left(f_0(c(t)), \frac{\partial f_0}{\partial x^i}(c(t))v_0^i \right).$$

This shows that $v(t) = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} c_s(t)$ is the integral curve of f_0^T with initial condition v_0.□

Tangent and cotangent lifts of the geodesic spray. Now let us consider the state manifold as the tangent bundle of the configuration manifold $M = TQ$. This allows us to take the drift vector field as the geodesic spray associated with the affine connection, $f_0 = Z$. Recall from equation (2.9) that

$$Z = v^i \frac{\partial}{\partial q^i} - \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i}.$$

Using coordinates $((q,v),(u,w))$ on TTQ, we find the tangent lift of the geodesic spray

$$\begin{aligned} Z^T = {} & v^i \frac{\partial}{\partial q^i} - \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i} + w^i \frac{\partial}{\partial u^i} \\ & - \left(\frac{\partial \Gamma^i_{jk}}{\partial q^l} v^j v^k u^l + \Gamma^i_{jk} w^j v^k + \Gamma^i_{kj} w^j v^k \right) \frac{\partial}{\partial w^i}. \end{aligned}$$

Following Theorem 3.5, we wish to get a handle on Z^T, related to classical notions from affine differential geometry. Let $c(t)$ be a geodesic on Q and let $c_s(t)$ be a family of geodesics with $c_0(t) = c(t)$. Define a vector field along $c(t)$ by

$$\xi(t) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} c_s(t) \in T_{c(t)}Q.$$

Kobayashi and Nomizu [1963b] show that this vector field satisfies the Jacobi equation

$$\nabla^2_{c'(t)}\xi(t) + R(\xi(t), c'(t))c'(t) + \nabla_{c'(t)}(T(\xi(t), c'(t))) = 0,$$

and so by definition $\xi(t)$ is a Jacobi field.

We want to find a relationship between variations of geodesics and variations of integral curves of Z. Now from Theorem 3.5, Z^T measures variations of integral curves of Z, and these solutions project to geodesics on Q. So we expect Z^T to measure variations of initial condition of geodesics. The Jacobi field measures variations of geodesics, so we also expect some relationship between Z^T and the Jacobi equation. Making precise such a relationship may give us insight into Z^{T^*}.

Our strategy will involve changing Z^T into a second-order vector field and using an ***Ehresmann connection*** to find a suitable way to split it, so that each part will satisfy certain conditions. This in turn will point us in the direction of what we want: a similar statement involving the cotangent lift of the geodesic spray Z^{T^*}.

The Jacobi equation is a second-order equation for a vector field along geodesics. But Z^T is not quite a second-order equation. However, if we define a diffeomorphism (canonical involution) $I_Q : TTQ \to TTQ$ by

$$I_Q((q,v),(u,w)) = ((q,u),(v,w)),$$

then $I_Q^* Z^T$ is a second-order vector field on TTQ. In coordinates it is

$$\begin{aligned} I_Q^* Z^T = u^i \frac{\partial}{\partial q^i} + w^i \frac{\partial}{\partial v^i} - \Gamma^i_{jk} u^j u^k \frac{\partial}{\partial u^i} \\ - \left(\frac{\partial \Gamma^i_{jk}}{\partial q^l} u^j u^k v^l + \Gamma^i_{jk} u^k w^j + \Gamma^i_{kj} u^k w^j \right) \frac{\partial}{\partial w^i}, \end{aligned}$$

which makes it clear that $I_Q^* Z^T$ is second-order.

We now try to gain some intuition about the geometry of the spaces $T_{v_q}TQ$ and $T^*_{v_q}TQ$ since we are investigating the maps

$$\begin{aligned} Z^T &: TTQ \to TTTQ, \\ Z^{T^*} &: T^*TQ \to TT^*TQ. \end{aligned}$$

Let us first try to get a picture of what $T_{v_q}TQ$ "looks" like. At $0_{q'} \in T_{q'}Q$, we may choose, what looks to be in Figure 1, a vertical component $V_{0_{q'}}TQ$, which is isomorphic to $T_{q'}Q$. A complement to this component, what we will refer to as the horizontal component $H_{0_{q'}}TQ$, can be taken as the space tangent to Q at the point q', which is simply $T_{q'}Q$. Thus we have split the space as

$$T_{0_{q'}}TQ \simeq T_{q'}Q \oplus T_{q'}Q.$$

If we take an arbitrary point $v_q \in T_qQ$, some extra care is required. The space T_qQ can still be taken as the vertical component, but now it is not clear what to choose for the horizontal component. So we always have a natural way to isolate the vertical component of a tangent space, that is

$$V_{v_q}TQ = \operatorname{span}\left\{ \frac{\partial}{\partial v^1}, \dots, \frac{\partial}{\partial v^n} \right\} \subset T_{v_q}TQ.$$

We want a complement to the vertical, and such a complement is called an ***Ehresmann connection***. To achieve this, we can use the affine connection ∇ to define the horizontal component as

$$H_{v_q}TQ = \operatorname{span}\left\{ \frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki}) v^k \frac{\partial}{\partial v^j} \right\}_{i \in \{1,\dots,n\}}.$$

This choice may be shown to be invariant under a change of coordinates, and so is well defined. It is also isomorphic to T_qQ because it can be projected to the tangent space at q.

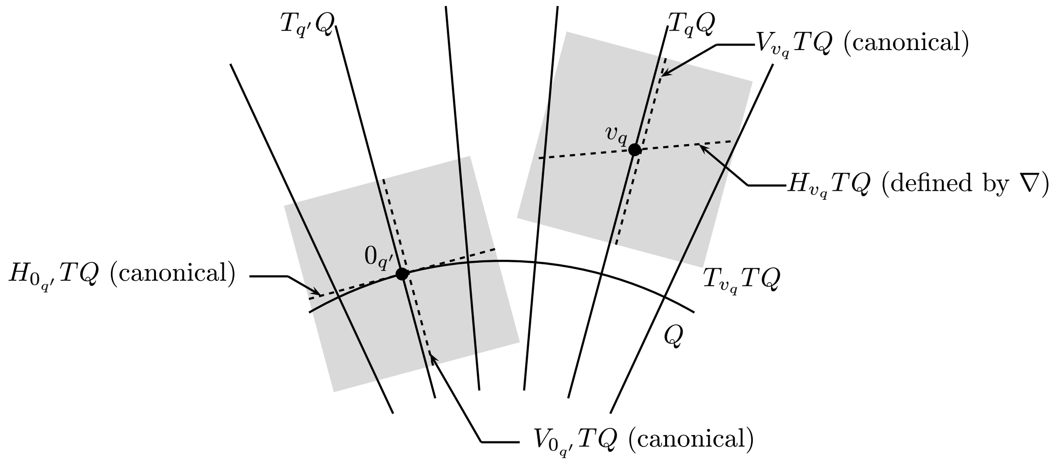

Figure 1: Splitting of the tangent space $T_{v_q}TQ$.

It is now evident that each of the components $V_{v_q}TQ$ and $H_{v_q}TQ$ are isomorphic to T_qQ, so we may write the splitting of our tangent space as

$$T_{v_q}TQ \simeq T_qQ \oplus T_qQ.$$

Since we used the canonical involution I_Q to make Z^T into a second-order vector field on TTQ, we can repeat the above process and use $I_Q^*Z^T$ to define a splitting of $T_{X_{v_q}}TTQ$ where $X_{v_q} \in T_{v_q}TQ$. Now, the vertical component is again naturally isomorphic to the tangent space $T_{v_q}TQ$. We then use another Ehresmann connection, obtained from $I_Q^*Z^T$, to find a complement $H_{v_q}TTQ$ that is also isomorphic to $T_{v_q}TQ$. This allows us to write

$$T_{X_{v_q}}TTQ \simeq T_{v_q}TQ \oplus T_{v_q}TQ,$$

and from the splitting found for $T_{v_q}TQ$, we have

$$T_{X_{v_q}}TTQ \simeq T_qQ \oplus T_qQ \oplus T_qQ \oplus T_qQ.$$

Therefore we can determine the form of $Z^T(X_{v_q})$ using this splitting, where $X_{v_q} \in T_{v_q}TQ$. Then we can use this representation of Z^T to obtain a relationship between solutions of the Jacobi equation and integral curves of Z^T.

Now for $X_{v_q} \in T_{v_q}TQ \simeq T_qQ \oplus T_qQ$, we can write

$$X_{v_q} = u_{v_q} \oplus w_{v_q},$$

where $u_{v_q}, w_{v_q} \in T_qQ$. With this in mind, we state the following theorem whose proof along the lines indicated above may be found in [Bullo and Lewis 2005, Chapter S4].

3.6 Theorem: *Let ∇ be an affine connection on Q with Z the corresponding geodesic spray. Let $c : I \to Q$ be a geodesic with $t \mapsto \sigma(t) \triangleq c'(t)$ the corresponding integral curve of Z. Let $a \in I$ and $u, w \in T_{c(a)}Q$ and define vector fields $U, W : I \to TQ$ along c by asking that $t \mapsto U(t) \oplus W(t) \in T_{c(t)}Q \oplus T_{c(t)}Q \simeq T_{\sigma(t)}TQ$ be the integral curve of Z^T with initial conditions $u \oplus w \in T_{c(a)}Q \oplus T_{c(a)}Q \simeq T_{\sigma(a)}TQ$. Then*

1. $W(t) = \nabla_{c'(t)}U(t) + \frac{1}{2}T(U(t), c'(t))$,
2. U *satisfies the Jacobi equation.*

The Jacobi equation is given in Definition 2.9,

$$\nabla^2_{c'(t)}\xi(t) + R(\xi(t), c'(t))c'(t) + \nabla_{c'(t)}(T(\xi(t), c'(t)) = 0,$$

where $\xi(t) \in TQ$ is a vector field along $c(t)$.

This gives us an understanding of the relationship between integral curves of Z^T and Jacobi fields. Now let us use the same approach to investigate the cotangent lift of the geodesic spray Z^{T^*}. For a vector $\Lambda_{v_q} \in T^*_{v_q}TQ$, we can define a splitting

$$T_{\Lambda_{v_q}}T^*TQ \simeq H_{\Lambda_{v_q}}T^*TQ \oplus V_{\Lambda_{v_q}}T^*TQ,$$

where $V_{\Lambda_{v_q}}T^*TQ$ is given naturally by $T^*_{v_q}TQ$ and $H_{\Lambda_{v_q}}T^*TQ$ is given by the Ehresmann connection defined by $I^*_Q Z^T$. Since the vector space $T_{v_q}TQ$ splits as

$$T_{v_q}TQ \simeq T_qQ \oplus T_qQ,$$

then its dual must also split as

$$V_{\Lambda_{v_q}}T^*TQ \simeq T^*_{v_q}TQ \simeq T^*_qQ \oplus T^*_qQ.$$

The horizontal component given by the Ehresmann connection turns out to be isomorphic to the tangent space to $T_{v_q}TQ$ itself. So we have

$$H_{\Lambda_{v_q}}T^*TQ \simeq T_{v_q}TQ \simeq T_qQ \oplus T_qQ.$$

Thus we have defined the splitting

$$T_{\Lambda_{v_q}}T^*TQ \simeq T_qQ \oplus T_qQ \oplus T^*_qQ \oplus T^*_qQ.$$

And as before, we can determine the form of $Z^{T^*}(\Lambda_{v_q})$ where $\Lambda_{v_q} \in T^*_{v_q}TQ$ using the above splitting.

Noting this, we can now write

$$\Lambda_{v_q} = \alpha_{v_q} \oplus \beta_{v_q},$$

where $\alpha_{v_q}, \beta_{v_q} \in T^*_qQ$. And this leads us to the following theorem, proven in [Bullo and Lewis 2005, Chapter S4].

3.7 Theorem: *Let ∇ be an affine connection on Q with Z the corresponding geodesic spray. Let $c : I \to Q$ be a geodesic with $t \mapsto \sigma(t) \triangleq c'(t)$ the corresponding integral curve of Z. Let $a \in I$ and $\alpha, \beta \in T^*_{c(a)}Q$ and define one-form fields $A, B : I \to T^*Q$ along c by asking that $t \mapsto A(t) \oplus B(t) \in T^*_{c(t)}Q \oplus T^*_{c(t)}Q \simeq T^*_{\sigma(t)}TQ$ be the integral curve of Z^{T^*} with initial conditions $\alpha \oplus \beta \in T^*_{c(a)}Q \oplus T^*_{c(a)}Q \simeq T^*_{\sigma(a)}TQ$. Then*

1. $A(t) = -\nabla_{c'(t)}B(t) + \frac{1}{2}T^*(B(t), c'(t))$,
2. B *satisfies the adjoint Jacobi equation.*

The adjoint Jacobi equation is given in Definition 2.10,

$$\nabla^2_{c'(t)}\alpha(t) + R^*(\alpha(t), c'(t))c'(t) - T^*(\nabla_{c'(t)}\alpha(t), c'(t)) = 0,$$

where $\alpha(t) \in T^*Q$ is a one-form field along $c(t)$.

3.3. The maximum principle for affine connection control systems. We now restate the maximum principle from Section 3.1, only we want to try to use the new concepts we have encountered. All definitions and theorems are from [Bullo and Lewis 2005, Chapter S4].

A ***control affine system*** is a triple $\Sigma = (M, \mathscr{F}, U)$ where M is an n-dimensional, separable, connected Hausdorff manifold, $\mathscr{F} = \{f_0, f_1, \ldots, f_m\}$ is a collection of C^∞ vector fields on M, and $U : M \to \mathbf{2}^{\mathbb{R}^m}$ assign a subset of $\mathbb{R}^m$ to each point $x \in M$, where $\mathbf{2}^A$ denotes the set of subsets of the set A. We denote $U_x \subset \mathbb{R}^m$ to be the image of $x \in M$ under U. If there is a subset $S \subset \mathbb{R}^m$ such that $U_x = S$ for each $x \in M$, then we say that U is *constant*.

Remember that we are dealing with the control system

$$\dot{x}(t) = f_0(x(t)) + u^a(t) f_a(x(t)).$$

We call a pair $\gamma = (u, c)$ a ***controlled trajectory*** for a control affine system $\Sigma = (M, \mathscr{F}, U)$ where $u : I \to \mathbb{R}^m$ is measurable and $c : I \to M$ is a solution of the control system, and denote the set of all control trajectories for Σ by $\mathrm{Ctraj}(\Sigma)$. If the interval I is compact, then we call $\gamma = (u, c)$ a ***controlled arc*** and denote the set of controlled arcs for Σ by $\mathrm{Carc}(\Sigma)$.

If the function $t \mapsto L(c(t), u(t))$ is locally integrable, then we will say that $\gamma = (u, c)$ is ***L-acceptable*** and denote $\mathrm{Ctraj}(\Sigma, L)$ as the subset of $\mathrm{Ctraj}(\Sigma)$ that contains only L-acceptable controlled trajectories. Similarly, $\mathrm{Carc}(\Sigma, L)$ is the subset of $\mathrm{Carc}(\Sigma)$ that contains only L-acceptable controlled arcs.

The functional that we are trying to minimize will be defined by

$$J^{\Sigma, L}(\gamma) = \int_a^b L(c(t), u(t))\, \mathrm{d}t,$$

where $\gamma = (u, c) \in \mathrm{Carc}(\Sigma, L)$ with u and c defined on $I = [a, b]$.

If S_0 and S_1 are two disjoint submanifolds of M, we define

$$\begin{aligned}\mathrm{Carc}(\Sigma, L, S_0, S_1) = \{\gamma = (u, c) \in \mathrm{Carc}(\Sigma, L) |\, & c(a) \in S_0 \text{ and } c(b) \in S_1 \\ & \text{where } u \text{ and } c \text{ are defined on } [a, b] \text{ for some } a, b \in \mathbb{R}\}.\end{aligned}$$

Now, if $a, b \in \mathbb{R}$ with $a < b$ are fixed, then we take

$$\begin{aligned}\mathrm{Carc}(\Sigma, L, S_0, S_1, [a, b]) = \{\gamma = (u, c) \in \mathrm{Carc}(\Sigma, L) |\, & \text{ where } u \text{ and } c \\ & \text{are defined on } [a, b] \text{ and } c(a) \in S_0 \text{ and } c(b) \in S_1\}.\end{aligned}$$

We can now formally state the problems of finding the optimal controlled trajectory connecting the two submanifolds.

3.8 Definition: Let $\Sigma = (M, \mathscr{F}, U)$ be a control affine system, let L be a cost for Σ, and let S_0 and S_1 be disjoint submanifolds of M.

1. A controlled arc $\gamma_* \in \mathrm{Carc}(\Sigma, L, S_0, S_1)$ is a ***solution of*** $\mathscr{P}(\Sigma, L, S_0, S_1)$ if $J^{\Sigma, L}(\gamma_*) \leq J^{\Sigma, L}(\gamma)$ for every $\gamma \in \mathrm{Carc}(\Sigma, L, S_0, S_1)$.

2. We say γ_* is a ***solution of*** $\mathscr{P}_{\mathrm{time}}(\Sigma, S_0, S_1)$ if it is a solution of $\mathscr{P}(\Sigma, 1, S_0, S_1)$.

3. A controlled arc $\gamma_* \in \mathrm{Carc}(\Sigma, L, S_0, S_1, [a, b])$ is a ***solution of*** $\mathscr{P}_{[a,b]}(\Sigma, L, S_0, S_1)$ if $J^{\Sigma, L}(\gamma_*) \leq J^{\Sigma, L}(\gamma)$ for every $\gamma \in \mathrm{Carc}(\Sigma, L, S_0, S_1, [a, b])$.

Given a real vector space V, its dual V^*, and a subset $S \subset V$, we denote

$$\mathrm{ann}(S) = \{\alpha \in V^* \,|\, \alpha(v) = 0 \text{ for all } V \in S\}.$$

We define the Hamiltonian $H^{\Sigma,L}$ on $\{(u, \alpha_x) \in U_x \times T^*M\}$ by

$$H^{\Sigma,L}(u, \alpha_x) = L(x, u) + \alpha_x \cdot (f_0(x) + u^a f_a(x)).$$

Notice that here we add L, whereas in the definition from Section 3.1, we subtracted it. This does not cause a problem since the one-form field will absorb the sign change.

We say that a one-form field $\chi : I \to T^*M$ along c is ***minimizing for*** $(\boldsymbol{\Sigma}, \boldsymbol{L})$ ***along*** $\boldsymbol{u}$ if

$$H^{\Sigma,L}(u(t), \chi(t)) \leq \inf_{\tilde{u} \in U_x} H^{\Sigma,L}(\tilde{u}, \chi(t)).$$

Using all this new terminology, we state another, more precise, version of the maximum principle.

3.9 Theorem: (Maximum Principle) *Let $\Sigma = (M, \mathscr{F}, U)$ be a control affine system with L a cost function for Σ, and let S_0 and S_1 be disjoint submanifolds of M. Suppose that $\gamma = (u, c) \in \mathrm{Carc}(\Sigma, L)$ is a solution of $\mathscr{P}_{[a,b]}(\Sigma, L, S_0, S_1)$ with u and c defined on $[a, b]$. Then there exists a one-form field $\chi : [a, b] \to T^*M$ along c and a constant $\chi_0 \in \{0, 1\}$ with the properties:*

1. *$\chi(a) \in \mathrm{ann}(T_{c(a)}S_0)$ and $\chi(b) \in \mathrm{ann}(T_{c(b)}S_1)$;*
2. *$\chi(t)$ satisfies the Hamiltonian equations for $H^{\Sigma,L}$ along u;*
3. *χ is minimizing for $(\Sigma, \chi_0 L)$ along u;*
4. *either $\chi_0 = 1$ or $\chi(a) \neq 0$.*

If U is constant then there exists a constant $C \in \mathbb{R}$ so that $H^{\Sigma,L}(u(t), \chi(t)) = C$ almost everywhere. If U is constant and if $\gamma = (u, c)$ is a solution of $\mathscr{P}(\Sigma, L, S_0, S_1)$, then we take $C = 0$.

This version of the maximum principle is a special case of a general maximum principle given by Sussmann [1998].

We now come to the main result of the paper, a maximum principle for affine connection control systems. However, since we are mostly interested in time-optimization, the theorem will only deal with the case when $L = 1$. For the general case, see [Bullo and Lewis 2005, Chapter S4]. But first we need some more notation.

An ***affine connection control system*** is a quadruple $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ where the configuration space Q is a smooth, finite-dimensional, separable, Hausdorff manifold, ∇ is a smooth affine connection on Q, $\mathscr{Y} = \{Y_1, \ldots, Y_m\}$ is a collection of smooth vector fields on Q, and $U : Q \to \mathbf{2}^{\mathbb{R}^m}$ is a map into the set of subsets of $\mathbb{R}^m$. Similarly to our previous description of the admissible controls, we denote $U_q \subset \mathbb{R}^m$, and say that U is ***constant*** if there is a subset $S \subset \mathbb{R}^m$ such that $U_q = U(q) = S$ for each $q \in Q$. Now we are considering the control systems characterized by the equation,

$$\nabla_{c'(t)} c'(t) = u^a(t) Y_a(c(t)). \tag{3.6}$$

In moving from the configuration space Q to the state space TQ, we find a control affine system $\Sigma = (TQ, \{Z, \mathrm{verlift}(Y_1), \dots, \mathrm{verlift}(Y_m)\}, U^T)$ associated with our affine connection control system Σ_{aff}, where $U^T(v_q) = U(q)$ and $\mathrm{verlift}(Y_a)(v_q) = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}(v_q + sY_a(q))$, so that if $Y_a = Y_a^i \frac{\partial}{\partial q^i}$ then $\mathrm{verlift}(Y_a) = Y_a^i \frac{\partial}{\partial v^i}$.

As we mentioned previously, for the problem of attempting to find time-optimal solutions, we will take the Lagrangian $L = 1$. So our problem is now to minimize

$$J^{\Sigma_{\mathrm{aff}}}(\gamma) = b - a.$$

We will not be dealing with the problems of fixed values for a and b, since this would imply that $J^{\Sigma_{\mathrm{aff}}}$ is constant for all trajectories and thus there would be nothing to minimize.

For $q_0, q_1 \in Q$, $v_{q_0} \in T_{q_0}Q$, and $v_{q_1} \in T_{q_1}Q$ we define

$$\begin{aligned}\mathrm{Carc}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1}) = \{\gamma = (u, c) \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}) |\, c'(a) = v_{q_0} \text{ and } c'(b) = v_{q_1} \\ \text{where } u \text{ and } c \text{ are defined on } [a,b] \text{ for some } a, b \in \mathbb{R}\}.\end{aligned}$$

This definition takes all controlled arcs that have fixed initial and final states, which restricts the configuration and the velocity. However, if we want to only restrict the configurations, we define

$$\begin{aligned}\mathrm{Carc}(\Sigma_{\mathrm{aff}}, q_0, q_1) = \{\gamma = (u, c) \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}) |\, c(a) = q_0 \text{ and } c(b) = q_1 \\ \text{where } u \text{ and } c \text{ are defined on } [a,b] \text{ for some } a, b \in \mathbb{R}\}.\end{aligned}$$

We define the time-optimal problem as follows:

3.10 Definition: Let $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be an affine connection control system, and let $L = 1$ be the cost function for Σ_{aff}, let $q_0, q_1 \in Q$, and let $v_{q_0} \in T_{q_0}Q$ and $T_{q_1} \in T_{q_1}Q$.

1. A controlled arc $\gamma_* \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1})$ is a ***solution of*** $\mathscr{P}_{\mathrm{time}}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1})$ if $J^{\Sigma_{\mathrm{aff}}}(\gamma_*) \le J^{\Sigma_{\mathrm{aff}}}(\gamma)$ for every $\gamma \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1})$.
2. A controlled arc $\gamma_* \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, q_0, q_1)$ is a ***solution of*** $\mathscr{P}_{\mathrm{time}}(\Sigma_{\mathrm{aff}}, q_0, q_1)$ if $J^{\Sigma_{\mathrm{aff}}}(\gamma_*) \le J^{\Sigma_{\mathrm{aff}}}(\gamma)$ for every $\gamma \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, q_0, q_1)$.

Before stating the next theorem, we refer back to Section 3.2, where we defined a splitting that allowed us to write the one-form field $\Lambda_{v_q} \in T^*_{v_q}TQ$ as $\alpha_{v_q} \oplus \beta_{v_q}$ for $\alpha_{v_q}, \beta_{v_q} \in T^*_qQ$. Using this fact, we define the ***time-optimal Hamiltonian*** $H^{\Sigma_{\mathrm{aff}}}_{\mathrm{time}}$ on $\{(u, \Lambda_{v_q}) \in U_q \times T^*TQ\}$ for an affine connection control system Σ_{aff} to be

$$H^{\Sigma_{\mathrm{aff}}}_{\mathrm{time}}(u, \alpha_{v_q} \oplus \beta_{v_q}) = 1 + \alpha_{v_q} \cdot v_q + u^a(\beta_{v_q} \cdot Y_a(q)). \tag{3.7}$$

A one-form field $\lambda : [a, b] \to T^*Q$ along c is ***minimizing for Σ_{aff} along u*** if

$$H^{\Sigma_{\mathrm{aff}}}_{\mathrm{time}}(u(t), \theta(t) \oplus \lambda(t)) \le \inf_{\tilde{u} \in U_q} H^{\Sigma_{\mathrm{aff}}}_{\mathrm{time}}(\tilde{u}, \theta(t) \oplus \lambda(t)).$$

Recall that in Theorem 3.7, it was found that the integral curve $A(t) \oplus B(t)$ of Z^{T^*} satisfies the equations

$$\begin{gathered}\nabla^2_{c'(t)} B(t) + R^*(B(t), c'(t))c'(t) - T^*(\nabla_{c'(t)} B(t), c'(t)) = 0, \\ A(t) = -\nabla_{c'(t)} B(t) + \frac{1}{2} T^*(B(t), c'(t)),\end{gathered}$$

the first being the adjoint Jacobi equation. We may now state the long awaited theorem for time-optimization on an affine connection control system. This is proven using Theorem 3.9 by Bullo and Lewis [2005, Chapter S4].

3.11 Theorem: (Maximum Principle for Time-Optimization of Affine Connection Control Systems) *Let $\Sigma_{\text{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be an affine connection control system with a cost function $L = 1$ for Σ_{aff}. Suppose that $\gamma = (u, c) \in \mathrm{Carc}(\Sigma_{\text{aff}})$ is a solution of $\mathscr{P}_{\text{time}}(\Sigma_{\text{aff}}, v_{q_0}, v_{q_1})$ with u and c defined on $[a, b]$. Then there exists a one-form field $\lambda : [a, b] \to T^*Q$ along c and a constant $\lambda_0 \in \{0, 1\}$ with the following properties:*

1. *for almost every $t \in [a, b]$ we have*
$$\nabla^2_{c'(t)}\lambda(t) + R^*(\lambda(t), c'(t))c'(t) - T^*(\nabla_{c'(t)}\lambda(t), c'(t)) = u^a(t)(\nabla Y_a)^*(\lambda(t)), \quad (3.8)$$
where $\nabla Y_a(X) = \nabla_X Y_a$;
2. *λ is minimizing for $(\Sigma_{\text{aff}}, \lambda_0)$ along u;*
3. *either $\lambda_0 = 1$ or $\theta(a) \oplus \lambda(a) \neq 0$, where*
$$\theta(t) = \frac{1}{2}T^*(\lambda(t), c'(t)) - \nabla_{c'(t)}\lambda(t), \quad t \in [a, b]. \quad (3.9)$$

If $\gamma = (u, c)$ is a solution of $\mathscr{P}_{\text{time}}(\Sigma_{\text{aff}}, q_0, q_1)$ then conditions 1–3 hold and, in addition, $\lambda(a) = 0$ and $\lambda(b) = 0$. If U is constant then λ may be chosen so that $H^{\Sigma_{\text{aff}}}_{\text{time}}(u(t), \theta(t) \oplus \lambda(t)) = 0$ almost everywhere.

3.4. Time-optimal control for affine connection control systems. As one could imagine, we need some kind of constraint on our controls, or else we would find that for time-optimization, "pushing harder" would achieve the end goal in zero time, which is unrealistic. Therefore let us assume that the norm of the controls with respect to a Riemannian metric g is bounded, $\|u^a(t)Y_a(c(t))\|_g \leq 1$. We would like to have a statement that tells us what the controls u^a are in this case. The following lemma achieves this result. If V is a vector space and V^* is its dual, we define $\sharp$ by stating that if $v \in V$ then $v^\sharp(u) = g(v, u)$ for every $u \in V$.

3.12 Lemma: *Given a vector space V with metric g and its dual space V^*, the problem of finding $v \in V$ that minimizes $\lambda(v)$ with the restriction that $\|v\| \leq 1$ where $\lambda \in V^*$, is solved by*
$$v_{\min} = -\frac{\lambda^\sharp}{\|\lambda\|_{g^{-1}}},$$
and the minimum value is given by $\lambda(v_{\min}) = -\|\lambda\|_{g^{-1}}$.

Proof: Choose an orthonormal basis $\{e_1, \ldots, e_n\}$ for V with
$$e_1 = \frac{\lambda^\sharp}{\|\lambda^\sharp\|_g}.$$

If $v = v^i e_i$, then

$$\lambda(v) = g(\lambda^\sharp, v) = \|\lambda^\sharp\|_g g(e_1, v) = \|\lambda^\sharp\|_g g(e_1, v^i e_i) = \|\lambda^\sharp\|_g v^1.$$

And since $-1 \leq v^i \leq 1$, then the minimum value of $\lambda(v)$ occurs when $v^1 = -1$, and so

$$v_{\min} = v^1 e_1 = -\frac{\lambda^\sharp}{\|\lambda^\sharp\|_g} = -\frac{\lambda^\sharp}{\|\lambda\|_{g^{-1}}}.$$

Knowing $v_{\min}$, we find that $\lambda(v_{\min}) = -\|\lambda\|_{g^{-1}}$.□

This lemma tells us that to minimize a one-form λ acting on a vector v whose norm is bounded by unity, the magnitude of the vector must be identically one, and its direction must be opposite to the image of the one form under $\sharp$.

We would like to apply this lemma to Theorem 3.11, the maximum principle for affine connection control systems. In this case, we will be trying to minimize the time-optimal Hamiltonian, equation (3.7):

$$H^{\Sigma_{\text{aff}}}_{\text{time}}(u, \alpha_{v_q} \oplus \beta_{v_q}) = 1 + \alpha_{v_q} \cdot v_q + u^a(\beta_{v_q} \cdot Y_a(q)).$$

Since the control does not appear in the first two terms $1 + \alpha_{v_q} \cdot v_q$, we are simply trying to find u^a such that

$$u^a(\beta_{v_q} \cdot Y_a(q)) = \min_{\|\tilde{u}^a Y_a(q)\|_g \leq 1} \tilde{u}^a(\beta_{v_q} \cdot Y_a(q)).$$

We are starting to see the how Lemma 3.12 can help us with such a problem.

Let $P_{\mathsf{Y}_q} : T_qQ \to T_qQ$ be the orthogonal projection onto the subspace spanned by our input vector fields, $\mathsf{Y}_q = \operatorname{span}\{Y_1(q), \ldots, Y_m(q)\}$. Using the subspace Y_q, we can split the tangent and cotangent spaces of the configuration manifold, so that

$$T_qQ \simeq \mathsf{Y}_q \oplus \mathsf{Y}_q^\perp \quad \text{and} \quad T_q^*Q \simeq \mathsf{Y}_q^* \oplus (\mathsf{Y}_q^\perp)^*.$$

Now we can write the one-form field $\beta_{v_q} \in T_q^*Q$ as

$$\beta_{v_q} = \beta^1_{v_q} + \beta^2_{v_q},$$

where $\beta^1_{v_q} = P^*_{\mathsf{Y}_q}(\beta_{v_q}) \in \mathsf{Y}_q^*$, and $\beta^2_{v_q} \in (\mathsf{Y}_q^\perp)^*$. The first component is the only one preserved when we apply the one-form β_{v_q} to the vector field $u^a Y_a(q)$. Thus it is quite obvious now how to apply Lemma 3.12.

Before stating the theorem for time-optimal control when the controls are bounded, we define the $(2,0)$-tensor on Q

$$h_{\mathsf{Y}_q}(\alpha_q, \beta_q) = g_{\mathsf{Y}_q}(g^\sharp(\alpha_q), g^\sharp(\beta_q)), \tag{3.10}$$

where g_{Y_q} is a $(0,2)$-tensor on Q given by

$$g_{\mathsf{Y}_q}(u_q, v_q) = g(P_{\mathsf{Y}_q}(u_q), P_{\mathsf{Y}_q}(v_q)). \tag{3.11}$$

This definition completes the notation that we need for the final result.

3.13 Theorem: *With the assumptions of Theorem 3.11 and the constraint that* $\|u^a(t)Y_a(c(t))\|_g \le 1$*, the time-optimal controls* $u^a(t)$ *are given by*

$$u^a(t)Y_a(c(t)) = -\frac{h^\sharp_{\mathsf{Y}_q}(\lambda(t))}{\|P^*_{\mathsf{Y}_q}(\lambda(t))\|_{g^{-1}}}. \tag{3.12}$$

provided that $h^\sharp_{\mathsf{Y}_q}(\lambda(t)) \neq 0$*.*

An extremal for which $h^\sharp_{\mathsf{Y}_q}(\lambda(t)) = 0$ along the entire extremal is called ***singular***. We will examine these in Section 5 when the planar rigid body is investigated.

4. Robotic leg

The first mechanical system that we will be investigating is known as the robotic leg. It consists of a retractable leg of variable length r with a mass m at one end. The other end is attached to a block, and both are free to rotate about this axis, as in Figure 2. We will measure the angular displacement of the block relative to the horizontal by θ and that of the leg extension by ψ. We assume that we can change (control) the length of the leg extension and the angle between the leg and block.

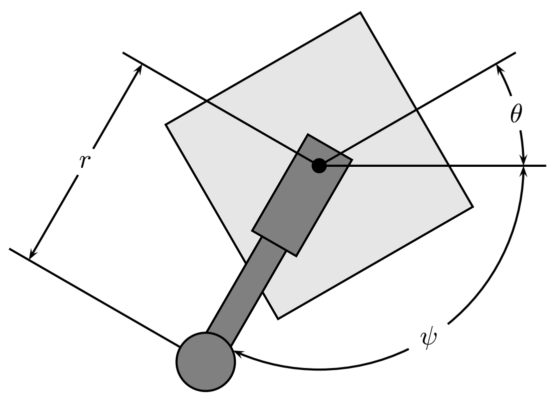

Figure 2: The robotic leg

4.1. Equations of motion. Taking the coordinates given above, (r, θ, ψ), the configuration manifold is $(0, \infty) \times \mathbb{T}^2$. The Riemannian metric for the robotic leg is

$$g = m\,\mathrm{d}r \otimes \mathrm{d}r + J\,\mathrm{d}\theta \otimes \mathrm{d}\theta + mr^2\,\mathrm{d}\psi \otimes \mathrm{d}\psi,$$

where m is the mass attached to the leg and J is the inertia of the block. To find the geodesics, we have to minimize the kinetic energy of the system, $\frac{1}{2}g(v_q, v_q)$, and this will describe its natural or unforced evolution in the configuration space. Thus the Lagrangian we will use to obtain these equations of motion is the kinetic energy,

$$L = \frac{1}{2}(m\dot{r}^2 + J\dot{\theta}^2 + mr^2\dot{\psi}^2).$$

The Euler-Lagrange equation then gives the uncontrolled motion of the system

$$\begin{aligned} m\ddot{r} - mr\dot{\psi}^2 &= 0, \\ J\ddot{\theta} &= 0, \\ mr^2\ddot{\psi} + 2mr\dot{r}\dot{\psi} &= 0. \end{aligned}$$

The input one-forms that determine how the controls are considered are $F^1 = \mathrm{d}r$ and $F^2 = \mathrm{d}\theta - \mathrm{d}\psi$. Therefore, the equations of controlled motion for this system are

$$\begin{aligned} m\ddot{r} - mr\dot{\psi}^2 &= u^1, \\ J\ddot{\theta} &= u^2, \\ mr^2\ddot{\psi} + 2mr\dot{r}\dot{\psi} &= -u^2, \end{aligned}$$

where u^1 is the leg-extension force and u^2 is the leg-body torque.

We could also find the equations of motion from the geodesic equation, $\nabla_{c'(t)}c'(t) = 0$. By taking $q^1 = r$, $q^2 = \theta$ and $q^3 = \psi$, the nonzero Christoffel symbols are

$$\Gamma^r_{\psi\psi} = -r \quad \text{and} \quad \Gamma^\psi_{r\psi} = \Gamma^\psi_{\psi r} = \frac{1}{r}.$$

The geodesics of uncontrolled motion are then obtained from equation (2.8):

$$\begin{aligned} \ddot{r} - r\dot{\psi}^2 &= 0, \\ \ddot{\theta} &= 0, \\ \ddot{\psi} + 2\frac{\dot{r}\dot{\psi}}{r} &= 0. \end{aligned}$$

The input vector fields Y_a are related to the input one-form fields F^a by the expression $Y_a = (F^a)^\sharp$, thus

$$Y_1 = \frac{1}{m}\frac{\partial}{\partial r} \quad \text{and} \quad Y_2 = \frac{1}{J}\frac{\partial}{\partial \theta} - \frac{1}{mr^2}\frac{\partial}{\partial \psi}.$$

Now equation (3.6), $\nabla_{c'(t)}c'(t) = u^a(t)Y_a(c(t))$ gives us the controlled equations of motion of the robotic leg to be

$$\begin{aligned} \ddot{r} - r\dot{\psi}^2 &= \frac{1}{m}u^1, \\ \ddot{\theta} &= \frac{1}{J}u^2, \\ \ddot{\psi} + 2\frac{\dot{r}\dot{\psi}}{r} &= -\frac{1}{mr^2}u^2, \end{aligned} \tag{4.1}$$

which are the same as those obtained from the Euler-Lagrange equation.

4.2. Application of the Hamiltonian equations. Since the geodesic equations are second-order, they can be written in the form of equation (3.5), the control equation

$$\dot{x}(t) = f_0(x(t)) + u^a(t)f_a(x(t)),$$

where we take $q(t) = (r(t), \theta(t), \psi(t))$ and $v(t) = (\dot{r}(t), \dot{\theta}(t), \dot{\psi}(t))$, so that $q \in Q$, the configuration manifold, and $v \in T_qQ$. So the state of our control system $x(t) = (q(t), v(t))$ changes according to equation (3.5) where

$$x = \begin{bmatrix} r \\ \theta \\ \psi \\ v_r \\ v_\theta \\ v_\psi \end{bmatrix}, f_0(x) = \begin{bmatrix} v_r \\ v_\theta \\ v_\psi \\ r v_\psi^2 \\ 0 \\ -2\frac{v_r v_\psi}{r} \end{bmatrix}, f_1(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{m} \\ 0 \\ 0 \end{bmatrix}, f_2(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{J} \\ -\frac{1}{mr^2} \end{bmatrix}.$$

Now, for $p = \begin{bmatrix} \alpha_r & \alpha_\theta & \alpha_\psi & \beta_r & \beta_\theta & \beta_\psi \end{bmatrix}$, a covector in $T^*_{(q,v)}TQ$, we may construct the control Hamiltonian, equation (3.4), remembering that for time-optimal control, the Lagrangian is $L = 1$.

$$\begin{aligned} H(x, p, u) &= p(f_0(x) + u^a f_a(x)) - p_0 L \\ &= \alpha_r v_r + \alpha_\theta v_\theta + \alpha_\psi v_\psi \\ &\quad + \frac{\beta_r}{m}(m r v_\psi^2 + u^1) + \frac{\beta_\theta}{J} u^2 - \frac{\beta_\psi}{mr^2}(2 m r v_r v_\psi + u^2) - p_0. \end{aligned}$$

The original maximum principle, Theorem 3.4, gives us the necessary condition that the Hamiltonian equations

$$\dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t), u(t)) \quad \text{and} \quad \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t), u(t))$$

must hold, giving us the system of differential equations

$$\begin{aligned} \dot{r} &= v_r, & \dot{\alpha}_r &= -\beta_r v_\psi^2 - 2\frac{\beta_\psi v_r v_\psi}{r^2} - 2\frac{\beta_\psi}{mr^3} u^2, \\ \dot{\theta} &= v_\theta, & \dot{\alpha}_\theta &= 0, \\ \dot{\psi} &= v_\psi, & \dot{\alpha}_\psi &= 0, \\ \dot{v}_r &= r v_\psi^2 + \frac{1}{m} u^1, & \dot{\beta}_r &= 2\frac{\beta_\psi v_\psi}{r} - \alpha_r, \\ \dot{v}_\theta &= \frac{1}{J} u^2, & \dot{\beta}_\theta &= -\alpha_\theta, \\ \dot{v}_\psi &= -2\frac{v_r v_\psi}{r} - \frac{1}{mr^2} u^2, & \dot{\beta}_\psi &= 2\frac{\beta_\psi v_r}{r} - 2\beta_r r v_\psi - \alpha_\psi. \end{aligned} \tag{4.2}$$

Of course, the obvious complexity of these equations suggest that we may not be able to find a closed form solution. However, here is where some numerical programming and simulations may offer insight.

4.3. Application of affine connection control systems. We now use the ideas we have gathered for time-optimization of affine connection control systems and apply it to the robotic leg example. First let us compute equation (3.8) from Theorem 3.11,

$$\begin{aligned} \nabla^2_{c'(t)}\lambda(t) + R^*(\lambda(t), c'(t))c'(t) - T^*(\nabla_{c'(t)}\lambda(t), c'(t)) & \\ = u^a(t)(\nabla Y_a)^*(\lambda(t)), & \end{aligned}$$

where $\lambda(t) = (\lambda_r(t), \lambda_\theta(t), \lambda_\psi(t))$ is a one-form field along $c(t) = (r(t), \theta(t), \psi(t))$, a trajectory in the configuration space Q.

As we mentioned earlier, the nonzero Christoffel symbols for the Riemannian metric are

$$\Gamma^r_{\psi\psi} = -r \quad \text{and} \quad \Gamma^\psi_{r\psi} = \Gamma^\psi_{\psi r} = \frac{1}{r}.$$

From the coordinate expressions in Definitions 2.6 and 2.7, we find that both the torsion tensor T and curvature tensor R are zero. And therefore the adjoint forms of each, T^* and R^*, are also zero. Using equation (2.7), we find the covariant derivative of $\lambda(t)$ along the trajectory to be

$$\nabla_{c'(t)}\lambda(t) = \left(\dot\lambda_r - \frac{\dot\psi\dot\lambda_\psi}{r}\right)\mathrm{d}r + \dot\lambda_\theta\,\mathrm{d}\theta + \left(\dot\lambda_\psi + r\dot\psi\lambda_r - \frac{\dot r\lambda_\psi}{r}\right)\mathrm{d}\psi.$$

Taking the second derivative gives us the right-hand side of equation (3.8),

$$\begin{aligned}
\nabla^2_{c'(t)}\lambda(t) &= \left(\ddot\lambda_r - \dot\psi^2\lambda_r - \frac{\ddot\psi\lambda_\psi + 2\dot\psi\dot\lambda_\psi}{r} + 2\frac{\dot r\dot\psi\lambda_\psi}{r^2}\right)\mathrm{d}r + \ddot\lambda_\theta\,\mathrm{d}\theta \\
&\quad + \left(\ddot\lambda_\psi + r\ddot\psi\lambda_r + 2r\dot\psi\dot\lambda_r - \dot\psi^2\lambda_\psi - \frac{\ddot r\lambda_\psi + 2\dot r\dot\lambda_\psi}{r} + 2\frac{\dot r^2\lambda_\psi}{r^2}\right)\mathrm{d}\psi \\
&= \left(\ddot\lambda_r - \dot\psi^2\lambda_r - 2\dot\psi\frac{r\dot\lambda_\psi - 2\dot r\lambda_\psi}{r^2}\right)\mathrm{d}r + \ddot\lambda_\theta\,\mathrm{d}\theta \\
&\quad + \Bigg(\ddot\lambda_\psi - 2\dot\psi(\dot\psi\lambda_\psi + \dot r\lambda_r - r\dot\lambda_r) \\
&\qquad\qquad - 2\dot r\frac{r\dot\lambda_\psi - \dot r\lambda_\psi}{r^2} - \frac{u^1\lambda_\psi + u^2\lambda_r}{mr}\Bigg)\mathrm{d}\psi,
\end{aligned}$$

where the second line comes from substituting $\ddot r$ and $\ddot\psi$ from the equations of motion, equations (4.1).

Now let us find the left-hand side of equation (3.8). In Theorem 3.11, we defined $\nabla Y_a(X) = \nabla_X Y_a$. From equation (2.4) we compute

$$\nabla Y_1(X) = \frac{1}{mr}X^3\frac{\partial}{\partial\psi} \implies \nabla Y_1 = \frac{1}{mr}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\nabla Y_2(X) = \frac{1}{mr}X^3\frac{\partial}{\partial r} + \frac{1}{mr^3}X^1\frac{\partial}{\partial\psi} \implies \nabla Y_2 = \frac{1}{mr}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ \frac{1}{r^2} & 0 & 0 \end{bmatrix},$$

where we are using matrix notation to denote ∇Y_a, so that $\nabla Y_a(X)$ is simply matrix multiplication of ∇Y_a and the vector $X = (X^1, X^2, X^3)$. And thus we find that

$$u^a(t)(\nabla Y_a)^*\lambda(t) = \frac{u^2\lambda_\psi}{mr^3}\,\mathrm{d}r + \frac{u^1\lambda_\psi + u^2\lambda_r}{mr}\,\mathrm{d}\psi.$$

So the maximum principle for affine connection control systems, Theorem 3.11 gives us the equations

$$\begin{aligned}
\ddot{\lambda}_r &= \dot{\psi}^2\lambda_r + 2\dot{\psi}\frac{r\dot{\lambda}_\psi - 2\dot{r}\lambda_\psi}{r^2},\\
\ddot{\lambda}_\theta &= 0,\\
\ddot{\lambda}_\psi &= 2\dot{\psi}(\dot{\psi}\lambda_\psi + \dot{r}\lambda_r - r\dot{\lambda}_r)\\
&\qquad + 2\dot{r}\frac{r\dot{\lambda}_\psi - \dot{r}\lambda_\psi}{r^2} + 2\frac{u^1\lambda_\psi + u^2\lambda_r}{mr},
\end{aligned}$$

which are equivalent to differentiating the Hamiltonian equations of equation (4.2) for $\dot{\beta}_r$, $\dot{\beta}_\theta$ and $\dot{\beta}_\psi$, after making some appropriate substitutions.

To use Theorem 3.13 to find the controls necessary for time-optimization in the case when we have the restriction $\|u^a(t)Y_a(c(t))\|_g \leq 1$, we first need a full orthonormal basis for the tangent bundle of the configuration manifold Q in terms of our input vector fields. Since Y_1 and Y_2 are orthogonal, we may choose any vector field that is orthogonal to both of these. Let us take $Y_3 = \frac{\partial}{\partial\theta} + \frac{\partial}{\partial\psi}$. Using the matrix notation we mentioned before, we find the orthonormal vector fields

$$\begin{aligned}
\tilde{Y}_1 &= \left(\frac{1}{\sqrt{m}}, 0, 0\right),\\
\tilde{Y}_2 &= \left(0, \sqrt{\frac{mr^2}{J(J+mr^2)}}, -\sqrt{\frac{J}{mr^2(J+mr^2)}}\right),\\
\tilde{Y}_3 &= \left(0, \frac{1}{\sqrt{J+mr^2}}, \frac{1}{\sqrt{J+mr^2}}\right).
\end{aligned}$$

This gives the orthogonal projection $P_{\mathsf{Y}_q} : T_qQ \to T_qQ$, onto the span of the input vector fields Y_1 and Y_2:

$$P_{\mathsf{Y}_q} = \frac{1}{J+mr^2}\begin{bmatrix} J+mr^2 & 0 & 0 \\ 0 & mr^2 & -mr^2 \\ 0 & -J & J \end{bmatrix}. \tag{4.3}$$

And therefore the $(0,2)$-tensor g_{Y_q} defined by equation (3.11) in the chosen coordinate system is

$$g_{\mathsf{Y}_q} = \frac{1}{J+mr^2}\begin{bmatrix} m(J+mr^2) & 0 & 0 \\ 0 & Jmr^2 & -Jmr^2 \\ 0 & -Jmr^2 & Jmr^2 \end{bmatrix}.$$

The numerator in equation (3.12) can be found by multiplying the matrix representing h_{Y_q} in equation (3.10) by the one-form field $\lambda(t) = (\lambda_r(t), \lambda_\theta(t), \lambda_\psi(t))$ to obtain

$$\begin{aligned}
h_{\mathsf{Y}_q} &= \frac{1}{J+mr^2}\begin{bmatrix} \frac{J+mr^2}{m} & 0 & 0 \\ 0 & \frac{mr^2}{J} & -1 \\ 0 & -1 & \frac{J}{mr^2} \end{bmatrix}\\
\implies h^\sharp_{\mathsf{Y}_q}(\lambda(t)) &= \frac{\lambda_r}{m}\frac{\partial}{\partial r} + \frac{\lambda_\theta mr^2 - \lambda_\psi J}{J(J+mr^2)}\frac{\partial}{\partial\theta} - \frac{\lambda_\theta mr^2 - \lambda_\psi J}{mr^2(J+mr^2)}\frac{\partial}{\partial\psi}.
\end{aligned}$$

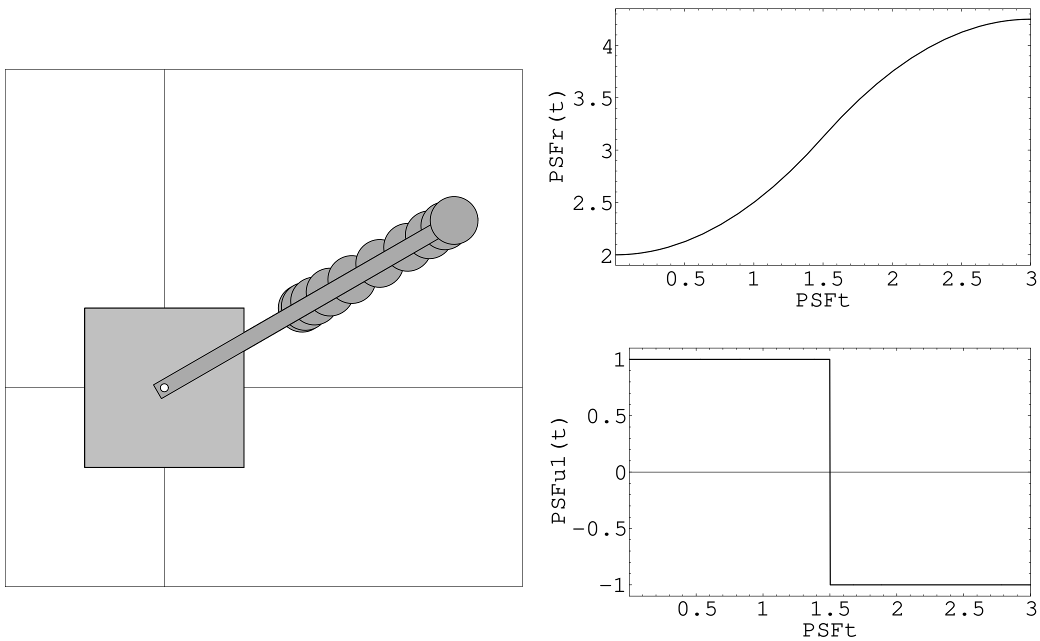

Figure 3: Example of time-optimal motion of the robotic leg when $m = 1$ and $J = 1$, and with initial conditions $q(0) = (2, 0, \frac{\pi}{6})$, $v(0) = (0, 0, 0)$, $\theta(0) = (-1, 0, 0)$ and $\lambda(0) = (-1.5, 0, 0)$. Plots of the leg extension length $r(t)$ and the control $u^1(t)$ are shown on the right.

Now, Lemma 3.12 tells us that the vector must be normalized, so we find the magnitude of the projection P_{Y_q} on the one-form field $\lambda(t)$,

$$\|P^*_{\mathsf{Y}_q}(\lambda(t))\|^2_{g^{-1}} = \frac{\lambda_r^2 Jr^2(J+mr^2) + \lambda_\theta^2 m^2 r^4 - 2\lambda_\theta\lambda_\psi Jmr^2 + \lambda_\psi^2 J^2}{Jmr^2(J+mr^2)},$$

where the norm is taken with respect to the inverse of the Riemannian metric g. And by applying equation (3.12), we find that our controls for time-optimal solutions along the one-form $\lambda(t)$ are

$$u^1 = -\frac{1}{\|P^*_{\mathsf{Y}_q}(\lambda(t))\|_{g^{-1}}}\lambda_r \quad \text{and} \quad u^2 = -\frac{1}{\|P^*_{\mathsf{Y}_q}(\lambda(t))\|_{g^{-1}}}\frac{\lambda_\theta mr^2 - \lambda_\psi J}{J+mr^2}.$$

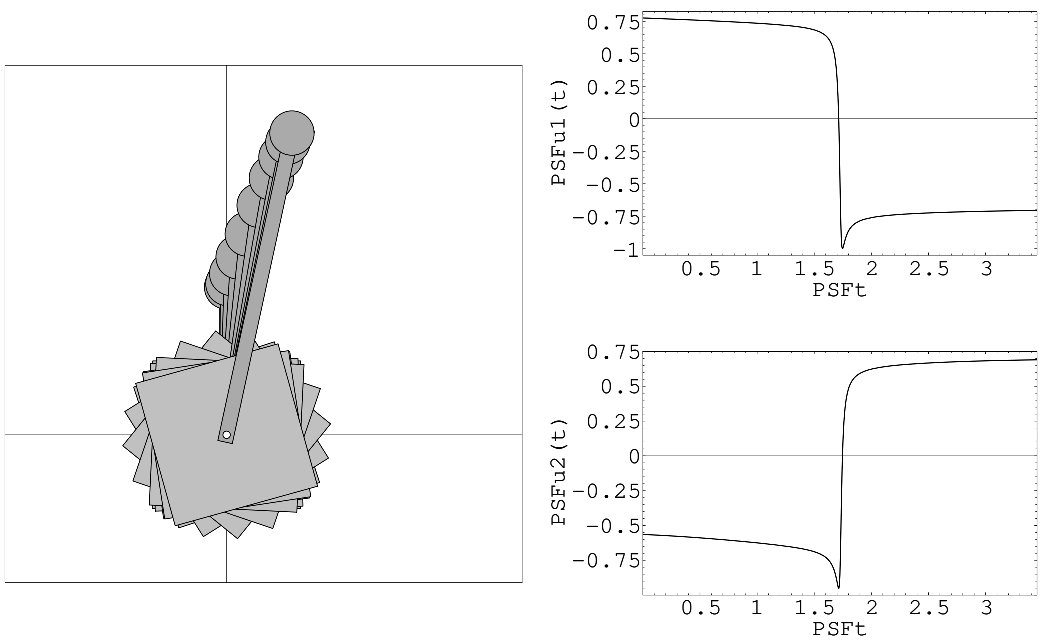

Figure 4: Example of time-optimal motion of the robotic leg when $m = 1$ and $J = 1$, and with initial conditions $q(0) = (2, 0, \frac{\pi}{2})$, $v(0) = (0, 0, 0)$, $\theta(0) = (-1, -1, -1)$ and $\lambda(0) = (-1.65, -2, -2)$. The controls are shown on the right.

5. Planar rigid body

The second mechanical system that we will explore is the planar rigid body. The body lies on a flat frictionless plane and we will assume that a force can be applied to the body at a point which is a distance h from the center of mass, as in Figure 5. Thus we will have an element Y_1 of the total force parallel to the line joining the point of application and the center of mass, and also an element Y_2 perpendicular to it. This system can be thought of as a simplified hovercraft.

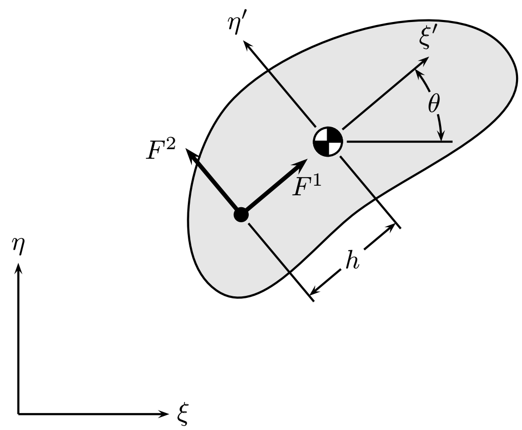

Figure 5: The planar rigid body

5.1. Equations of motion. We will take the Cartesian coordinate system $(\xi, \eta) \in \mathbb{R}^2$ to describe the position of the body, and let θ denote the angle that the line joining the point of application of the force and the center of mass makes with the horizontal. So the configuration manifold for these coordinates is $\mathbb{R}^2 \times \mathbb{S}^1$. Using these coordinates (ξ, η, θ), the Riemannian metric for the planar rigid body is

$$g = m\,\mathrm{d}\xi \otimes \mathrm{d}\xi + m\,\mathrm{d}\eta \otimes \mathrm{d}\eta + J\,\mathrm{d}\theta \otimes \mathrm{d}\theta,$$

where m is the mass of the body and J is the inertia. By the same reasoning for the robotic leg, we take the Lagrangian to be the kinetic energy $\frac{1}{2}g(v_q, v_q)$ of the system,

$$L = \frac{1}{2}(m\dot{\xi}^2 + m\dot{\eta}^2 + J\dot{\theta}^2).$$

Euler's necessary condition, Theorem 2.2, tells us that the geodesic equations of the system are $m\ddot{\xi} = 0$, $m\ddot{\eta} = 0$, and $J\ddot{\theta} = 0$. The input one-forms that give us the components of force and the moment of inertia applied to the system from our controls are $F^1 = \cos\theta\,\mathrm{d}\xi + \sin\theta\,\mathrm{d}\eta$ and $F^2 = -\sin\theta\,\mathrm{d}\xi + \cos\theta\,\mathrm{d}\eta - h\,\mathrm{d}\theta$. Therefore we find that the equations of controlled motion for this system are

$$\begin{aligned} m\ddot{\xi} &= u^1\cos\theta - u^2\sin\theta, \\ m\ddot{\eta} &= u^1\sin\theta + u^2\cos\theta, \\ J\ddot{\theta} &= -hu^2, \end{aligned}$$

where u^1 is the component of the force along the line joining its point of application and the center of mass, and u^2 is the component of the force perpendicular to this line.

To find the equations of motion from the geodesic equation (2.8), we note that all of the Christoffel symbols are zero since the metric g is constant. This tells us that the uncontrolled equations are simply the second derivatives of each coordinate $q^1 = \xi$, $q^2 = \eta$, and $q^3 = \theta$ equal to zero, as we found earlier. We compute the input vector fields, again from the relation $Y_a = (F^a)^\sharp$, to be

$$\begin{aligned} Y_1 &= \frac{1}{m}\cos\theta\frac{\partial}{\partial\xi} + \frac{1}{m}\sin\theta\frac{\partial}{\partial\eta}, \\ Y_2 &= -\frac{1}{m}\sin\theta\frac{\partial}{\partial\xi} + \frac{1}{m}\cos\theta\frac{\partial}{\partial\eta} - \frac{h}{J}\frac{\partial}{\partial\theta}. \end{aligned}$$

And so equation (3.6), $\nabla_{c'(t)}c'(t) = u^a(t)Y_a(c(t))$ gives the same equations of controlled motion of the planar rigid body as the Euler-Lagrange equation gave us,

$$\begin{aligned} \ddot{\xi} &= \frac{1}{m}u^1\cos\theta - \frac{1}{m}u^2\sin\theta, \\ \ddot{\eta} &= \frac{1}{m}u^1\sin\theta + \frac{1}{m}u^2\cos\theta, \\ \ddot{\theta} &= -\frac{h}{J}u^2. \end{aligned} \tag{5.1}$$

5.2. Application of the Hamiltonian equations. To find the Hamiltonian equations, we repeat the same procedure used on the robotic leg in Section 4.2. Letting $q(t) = (\xi(t), \eta(t), \theta(t))$ and $v(t) = (\dot{\xi}(t), \dot{\eta}(t), \dot{\theta}(t))$ so that we can again write the second-order non-linear equations of motion as equation (3.5)

$$\dot{x}(t) = f_0(x(t)) + u^a(t)f_a(x(t)),$$

where $x(t) = (q(t), v(t))$ is the state of the system. To achieve this, we take the assignments:

$$x = \begin{bmatrix} \xi \\ \eta \\ \theta \\ v_\xi \\ v_\eta \\ v_\theta \end{bmatrix}, f_0(x) = \begin{bmatrix} v_\xi \\ v_\eta \\ v_\theta \\ 0 \\ 0 \\ 0 \end{bmatrix}, f_1(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{m}\cos\theta \\ \frac{1}{m}\sin\theta \\ 0 \end{bmatrix}, f_2(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\frac{1}{m}\sin\theta \\ \frac{1}{m}\cos\theta \\ -\frac{h}{J} \end{bmatrix}.$$

From equation (3.4), the Hamiltonian for time-optimal control becomes

$$\begin{aligned} H(x,p,u) = {} & \alpha_\xi v_\xi + \alpha_\eta v_\eta + \alpha_\theta v_\theta + \\ & \frac{\beta_\xi}{m}(u^1\cos\theta - u^2\sin\theta) + \frac{\beta_\eta}{m}(u^1\sin\theta + u^2\cos\theta) - \frac{\beta_\theta}{J}hu^2 - p_0, \end{aligned}$$

where $p = \begin{bmatrix} \alpha_\xi & \alpha_\eta & \alpha_\theta & \beta_\xi & \beta_\eta & \beta_\theta \end{bmatrix} \in T^*_{(q,v)}TQ$ is a one-form field along a solution c. And therefore the Hamiltonian equations from the maximum principle,

$$\dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t), u(t)) \quad \text{and} \quad \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t), u(t)),$$

tell us that the time-optimal trajectories must satisfy

$$
\begin{aligned}
\dot{\xi} &= v_\xi, & \dot{\alpha}_\xi &= 0,\\
\dot{\eta} &= v_\eta, & \dot{\alpha}_\eta &= 0,\\
\dot{\theta} &= v_\theta, & \dot{\alpha}_\theta &= \frac{1}{m}(\beta_\xi u^1 + \beta_\eta u^2)\sin\theta\\
\dot{v}_\xi &= \frac{1}{m}(u^1\cos\theta - u^2\sin\theta), & &\quad + \frac{1}{m}(\beta_\xi u^2 - \beta_\eta u^1)\cos\theta,\\
\dot{v}_\eta &= \frac{1}{m}(u^1\sin\theta + u^2\cos\theta), & \dot{\beta}_\xi &= -\alpha_\xi,\\
\dot{v}_\theta &= -\frac{h}{J}u^2, & \dot{\beta}_\eta &= -\alpha_\eta,\\
& & \dot{\beta}_\theta &= -\alpha_\theta.
\end{aligned}
\tag{5.2}
$$

From this system, we see immediately that β_ξ and β_η are linear functions of time:

$$
\begin{aligned}
\beta_\xi &= k_1 - c_1 t,\\
\beta_\eta &= k_2 - c_2 t,
\end{aligned}
\tag{5.3}
$$

where $c_1 = \alpha_\xi$, $c_2 = \alpha_\eta$, k_1, and k_2 are constants determined by the boundary conditions. This will prove useful in Section 5.4.

5.3. Application of affine connection control systems. To apply the results of time-optimal control of affine connection control systems, Theorems 3.11 and 3.13, we first recall that all the Christoffel symbols are zero in the Levi-Civita connection chosen. From this fact, it is apparent that the adjoint torsion tensor T^* and adjoint curvature tensor R^* are also zero. So to find equation (3.8),

$$\nabla^2_{c'(t)}\lambda(t) = u^a(t)(\nabla Y_a)^*\lambda(t),$$

where $\lambda(t) = (\lambda_\xi(t), \lambda_\eta(t), \lambda_\theta(t))$, we first note that

$$\nabla^2_{c'(t)}\lambda(t) = \ddot{\lambda}_\xi\,\mathrm{d}\xi + \ddot{\lambda}_\eta\,\mathrm{d}\eta + \ddot{\lambda}_\theta\,\mathrm{d}\theta.$$

From the input vector fields

$$
\begin{aligned}
Y_1 &= \frac{1}{m}\cos\theta\frac{\partial}{\partial\xi} + \frac{1}{m}\sin\theta\frac{\partial}{\partial\eta},\\
Y_2 &= -\frac{1}{m}\sin\theta\frac{\partial}{\partial\xi} + \frac{1}{m}\cos\theta\frac{\partial}{\partial\eta} - \frac{h}{J}\frac{\partial}{\partial\theta},
\end{aligned}
$$

given in section 5.1, we find the terms $\nabla Y_a(X) = \nabla_X Y_a$ from equation (2.4)

$$
\begin{aligned}
\nabla Y_1(X) &= -\frac{1}{m}\sin\theta\, X^3\frac{\partial}{\partial\xi} + \frac{1}{m}\cos\theta\, X^3\frac{\partial}{\partial\eta}\\
\implies \nabla Y_1 &= \frac{1}{m}\begin{bmatrix} 0 & 0 & -\sin\theta\\ 0 & 0 & \cos\theta\\ 0 & 0 & 0\end{bmatrix},\\
\nabla Y_2(X) &= -\frac{1}{m}\cos\theta\, X^3\frac{\partial}{\partial\xi} - \frac{1}{m}\sin\theta\, X^3\frac{\partial}{\partial\eta}\\
\implies \nabla Y_2 &= \frac{1}{m}\begin{bmatrix} 0 & 0 & -\cos\theta\\ 0 & 0 & -\sin\theta\\ 0 & 0 & 0.\end{bmatrix}.
\end{aligned}
$$

Therefore the left-hand side of equation (3.8) is

$$u^a(t)(\nabla Y_a)^*\lambda(t) = -\frac{1}{m}\big((u^1\lambda_\xi + u^2\lambda_\eta)\sin\theta \\ + (u^2\lambda_\xi - u^1\lambda_\eta)\cos\theta\big)\,\mathrm{d}\theta,$$

and thus we find the same equations that are given after differentiating $\dot\beta_\xi$, $\dot\beta_\eta$, and $\dot\beta_\theta$ from the Hamiltonian equations, equations (5.2):

$$\begin{aligned}
\ddot\lambda_\xi &= 0,\\
\ddot\lambda_\eta &= 0,\\
\ddot\lambda_\theta &= -\frac{1}{m}(\lambda_\xi u^1 + \lambda_\eta u^2)\sin\theta - \frac{1}{m}(\lambda_\xi u^2 - \lambda_\eta u^1)\cos\theta.
\end{aligned} \tag{5.4}$$

In applying Theorem 3.13 to find the time-optimal controls for the planar rigid body example when $\|u^a(t)Y_a(c(t))\|_g \le 1$, we first try to simplify the calculations by introducing the rotation matrix

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We then choose the vector field $Y_3 = \sin\theta\frac{\partial}{\partial\xi} - \cos\theta\frac{\partial}{\partial\eta} - \frac{1}{h}\frac{\partial}{\partial\theta}$ for a complement to the orthogonal pair Y_1 and Y_2, and so we compute the orthonormal vector fields

$$\begin{aligned}
\tilde Y_1 &= \frac{1}{\sqrt{m}}R(\theta)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\\
\tilde Y_2 &= \sqrt{\frac{J}{m(J+mh^2)}}R(\theta)\begin{bmatrix} 0 \\ 1 \\ -\frac{mh}{J} \end{bmatrix},\\
\tilde Y_3 &= \frac{h}{\sqrt{J+mh^2}}R(\theta)\begin{bmatrix} 0 \\ -1 \\ -\frac{1}{h} \end{bmatrix}.
\end{aligned}$$

This allows us to represent the orthogonal projection $P_{\mathsf{Y}_q} : T_qQ \to T_qQ$ by the matrix

$$P_{\mathsf{Y}_q} = \frac{1}{J+mh^2}R(\theta)\begin{bmatrix} J+mh^2 & 0 & 0 \\ 0 & J & -Jh \\ 0 & -mh & mh^2 \end{bmatrix}R^{-1}(\theta). \tag{5.5}$$

Now we use the fact that the metric g commutes with the rotation matrix $R(\theta)$ and also that $R^{-1}(\theta) = R^*(\theta)$. This lets us write the identity $R^{-1}(\theta)gR(\theta) = g$ and so we find that

$$g_{\mathsf{Y}_q} = \frac{1}{J+mh^2}R(\theta)\begin{bmatrix} m(J+mh^2) & 0 & 0 \\ 0 & Jm & -Jmh \\ 0 & -Jmh & Jmh^2 \end{bmatrix}R^{-1}(\theta).$$

And then we compute $h^{\sharp}_{\Upsilon_q}(\lambda(t))$ from equation (3.10), first noting that

$$h_{\Upsilon_q} = \frac{1}{J+mh^2} R(\theta) \begin{bmatrix} \frac{J+mh^2}{m} & 0 & 0 \\ 0 & \frac{J}{m} & -h \\ 0 & -h & \frac{mh^2}{J} \end{bmatrix} R^{-1}(\theta).$$

So the numerator in equation (3.12) is

$$\begin{aligned} h^{\sharp}_{\Upsilon_q}(\lambda(t)) = \frac{1}{J+mh^2}\Big(\lambda_\xi\big(\tfrac{J}{m} + h^2\cos^2\theta\big) + \lambda_\eta h^2 \sin\theta\cos\theta + \lambda_\theta h \sin\theta\Big)\frac{\partial}{\partial \xi} \\ + \frac{1}{J+mh^2}\Big(\lambda_\xi h^2 \sin\theta\cos\theta + \lambda_\eta\big(\tfrac{J}{m} + h^2\sin^2\theta\big) - \lambda_\theta h\cos\theta\Big)\frac{\partial}{\partial y} \\ + \frac{1}{J+mh^2}\Big(\lambda_\xi h\sin\theta - \lambda_\eta h\cos\theta + \lambda_\theta \frac{mh^2}{J}\Big)\frac{\partial}{\partial\theta}. \end{aligned}$$

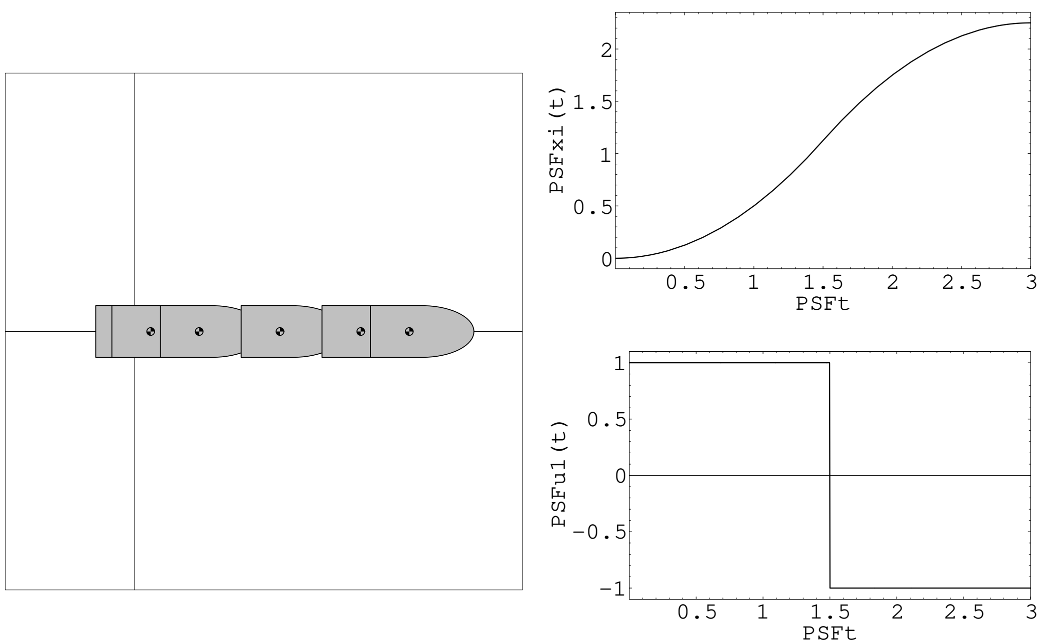

Figure 6: Example of time-optimal motion of the planar rigid body when $m = 1$, $J = 1$, and $h = 1$, and with initial conditions $q(0) = (0,0,0)$, $v(0) = (0,0,0)$, $\theta(0) = (-1,0,0)$ and $\lambda(0) = (-1.5,0,0)$. Plots of the coordinate $\xi(t)$ and the control $u^1(t)$ are shown on the right.

Now, from equation (5.5) we find the square of the norm of $P^*_{\mathsf{Y}_q}(\lambda(t))$:

$$\|P^*_{\mathsf{Y}_q}(\lambda(t))\|^2_{g^{-1}} = \frac{1}{m(J+mh^2)}\Big(\lambda_\xi^2(J+mh^2\cos^2\theta) + 2\lambda_\xi\lambda_\eta mh^2\sin\theta\cos\theta + \lambda_\eta^2(J+mh^2\sin^2\theta) - 2\lambda_\eta\lambda_\theta mh\cos\theta + \lambda_\theta^2\frac{m^2h^2}{J} + 2\lambda_\xi\lambda_\theta mh\sin\theta\Big).$$

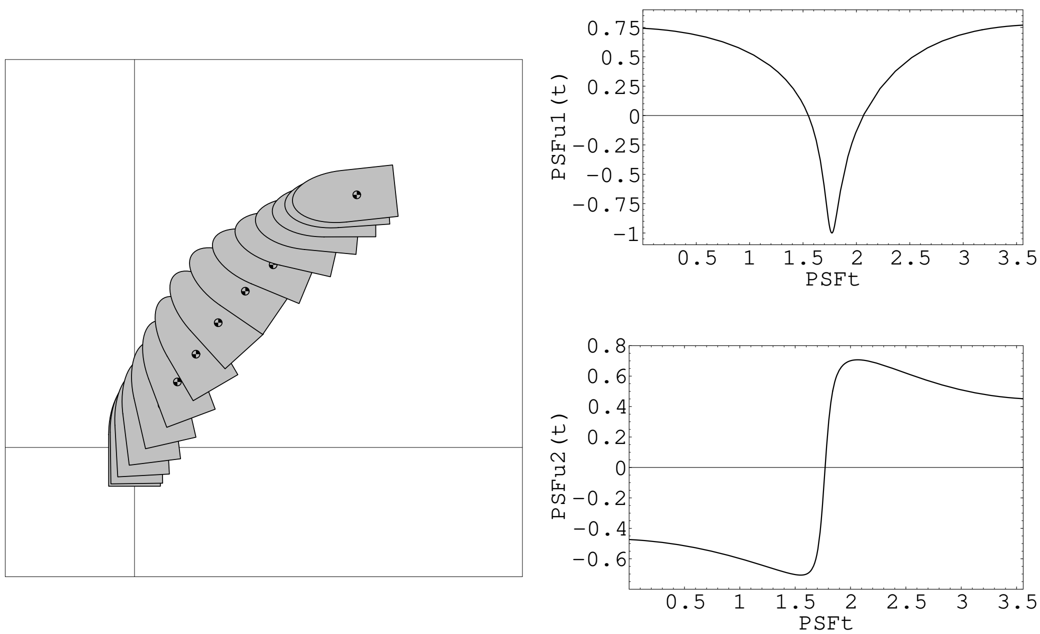

Figure 7: Example of time-optimal motion of the planar rigid body when $m=1$, $J=1$, and $h=1$, and with initial conditions $q(0)=(0,0,\frac{\pi}{2})$, $v(0)=(0,0,0)$, $\theta(0)=(-1.25,-1,0.325)$ and $\lambda(0)=(-2.55,-2,0)$. The controls are shown on the right.

Theorem 3.13 then tells us that when the norm of

$$u^1Y_1(q)+u^2Y_2(q) = \frac{1}{m}(u^1\cos\theta - u^2\sin\theta)\frac{\partial}{\partial\xi} + \frac{1}{m}(u^1\sin\theta+u^2\cos\theta)\frac{\partial}{\partial\eta} - u^2\frac{h}{J}\frac{\partial}{\partial\theta}$$

is bounded by one, we can write our controls for time-optimization in terms of the one-form $\lambda(t)$ along a solution,

$$u^1 = -\frac{1}{\|P^*_{\mathsf{Y}_q}(\lambda(t))\|_{g^{-1}}}\Big(\lambda_\xi\cos\theta + \lambda_\eta\sin\theta\Big),$$
$$u^2 = \frac{1}{\|P^*_{\mathsf{Y}_q}(\lambda(t))\|_{g^{-1}}}\frac{\lambda_\xi J\sin\theta - \lambda_\eta J\cos\theta + \lambda_\theta mh}{J+mh^2}.$$

5.4. Singular extremals. We now look at the special case when $h^\sharp_{\mathsf{Y}_q}(\lambda(t)) = 0$. Extremals with this property are called ***singular***, and occur when the maximum principle does not allow us to determine the optimal controls. But for the planar rigid body, this condition does give us the following equations:

$$\begin{aligned}
\lambda_\xi\left(\frac{J}{m} + h^2\cos^2\theta\right) + \lambda_\eta h^2\sin\theta\cos\theta + \lambda_\theta h\sin\theta &= 0,\\
\lambda_\xi h^2\sin\theta\cos\theta + \lambda_\eta\left(\frac{J}{m} + h^2\sin^2\theta\right) - \lambda_\theta h\cos\theta &= 0,\\
\lambda_\xi h\sin\theta - \lambda_\eta h\cos\theta + \lambda_\theta\frac{mh^2}{J} &= 0.
\end{aligned} \tag{5.6}$$

The first two of these give a relation between θ, λ_ξ, and λ_η,

$$\lambda_\xi\cos\theta + \lambda_\eta\sin\theta = 0. \tag{5.7}$$

Using this relation and the third in equations (5.6), we find two more that will prove quite useful,

$$\lambda_\xi + \frac{mh}{J}\lambda_\theta\sin\theta = 0 \quad\text{and}\quad \lambda_\eta - \frac{mh}{J}\lambda_\theta\cos\theta = 0. \tag{5.8}$$

Using the identity $\sin^2\theta + \cos^2\theta = 1$, we obtain the fact that

$$\lambda_\theta^2 = \left(\frac{J}{mh}\right)^2(\lambda_\xi^2 + \lambda_\eta^2).$$

Now, in Section 5.2, we found β_ξ and β_η were linear functions of time. From these, equations (5.3), we have an explicit form of λ_θ,

$$\lambda_\theta^2 = \left(\frac{J}{mh}\right)^2\big((k_1 - c_1t)^2 + (k_2 - c_2t)^2\big).$$

Knowing the one-forms λ_ξ, λ_η, and λ_θ in terms of the independent variable t allows us to find θ in a closed form from equations (5.8) since

$$\theta = \arctan\left(\frac{-\lambda_\xi}{\lambda_\eta}\right).$$

Then we may use $\ddot\theta = -\frac{h}{J}u^2$, one of the equations of motion, equations (5.1), to find an explicit formula for u^2 in terms of t. Now we can take the third of equations (5.4), $\ddot\lambda_\theta = -\frac{1}{m}(\lambda_\xi u^1 + \lambda_\eta u^2)\sin\theta - \frac{1}{m}(\lambda_\xi u^2 - \lambda_\eta u^1)\cos\theta$, and solve for u^1. It is then possible to find ξ and η by integrating twice the other equations of motion,

$$\begin{aligned}
\ddot\xi &= \frac{1}{m}(u^1\cos\theta - u^2\sin\theta),\\
\ddot\eta &= \frac{1}{m}(u^1\sin\theta + u^2\cos\theta),
\end{aligned}$$

since we know all the variables on the left-hand sides. So for the planar rigid body example, when $h^\sharp_{\mathsf{Y}_q}(\lambda(t)) = 0$ we can obtain closed form solutions for $q(t)$ and $\lambda(t)$, and also the controls needed for time-optimization.

Using the method described above, we find that the trajectories of singular extremals are governed by the equations

$$\begin{aligned}
\xi(t) &= -\frac{J(k_2 - c_2 t)}{mh\sqrt{(k_1 - c_1 t)^2 + (k_2 - c_2 t)^2}} + \xi_0 t + \xi_1, \\
\eta(t) &= \frac{J(k_1 - c_1 t)}{mh\sqrt{(k_1 - c_1 t)^2 + (k_2 - c_2 t)^2}} + \eta_0 t + \eta_1, \\
\theta(t) &= \arctan\left(\frac{-k_1 + c_1 t}{k_2 - c_2 t}\right),
\end{aligned} \tag{5.9}$$

where ξ_0, ξ_1, η_0, and η_1 are the integration constants that occur from integrating $\ddot{\xi}(t)$ and $\ddot{\eta}(t)$ twice. The controls may also be found explicitly, given by

$$\begin{aligned}
u^1(t) &= \frac{J(k_1 c_2 - k_2 c_1)^2}{h((k_1 - c_1 t)^2 + (k_2 - c_2 t)^2)^2}, \\
u^2(t) &= -\frac{2J(k_1 c_2 - k_2 c_1)((c_1^2 + c_2^2)t - (c_1 k_1 + c_2 k_2))}{h((k_1 - c_1 t)^2 + (k_2 - c_2 t)^2)^2}.
\end{aligned}$$

The linear functions $\xi_0 t + \xi_1$ and $\eta_0 t + \eta_1$ simply translate the body along a straight line as t increases. For the moment, let us ignore these translations and take the constants ξ_0, ξ_1, η_0, and η_1 as zero, to allow us to examine the nature of the singular extremals more easily.

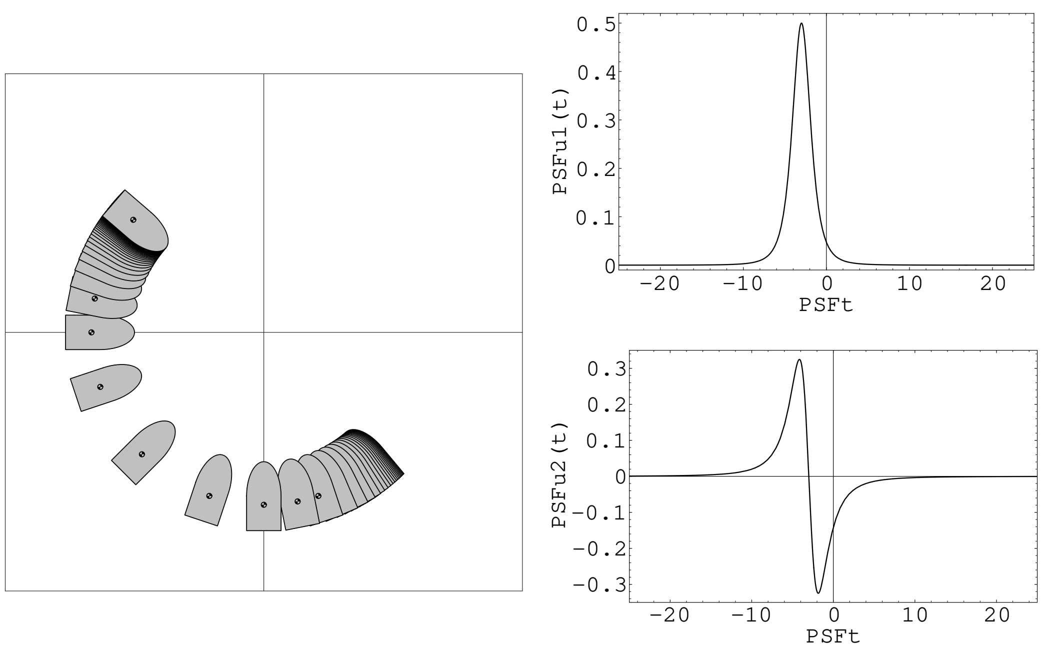

Figure 8: Example of the motion of singular extremals without linear translation when $m = 1$, $J = 2$, and $h = 1$, and with constants $c_1 = -1$, $c_2 = -1$, $k_1 = 1$, and $k_2 = 5$. The controls are shown on the right.

When the singular extremals are not translated, their equations are:

$$\begin{aligned} \xi(t) &= -\frac{J(k_2 - c_2 t)}{mh\sqrt{(k_1 - c_1 t)^2 + (k_2 - c_2 t)^2}}, \\ \eta(t) &= \frac{J(k_1 - c_1 t)}{mh\sqrt{(k_1 - c_1 t)^2 + (k_2 - c_2 t)^2}}, \\ \theta(t) &= \arctan\left(\frac{-k_1 + c_1 t}{k_2 - c_2 t}\right). \end{aligned} \tag{5.10}$$

To characterize the path of such extremals, we find by a direct calculation that

$$\xi^2(t) + \eta^2(t) = \left(\frac{J}{mh}\right)^2 \quad \text{and} \quad \theta(t) = \arctan\left(\frac{\eta(t)}{\xi(t)}\right) + \pi.$$

Thus when the constants of integration are zero, all motions of the planar rigid body are along circles in the (ξ, η)-plane of radius $\frac{J}{mh}$ while the line joining the point of application of the force and the center of mass points directly toward the center of this circle, as in Figure 8.

We now return to the more general class of singular extremals given by equations (5.9). Let us consider the *unforced extremals* which occur when $u^1(t) = 0$ and $u^2(t) = 0$. One may verify that this holds when

$$k_1 c_2 = k_2 c_1. \tag{5.11}$$

By differentiating the first two in equations (5.10), we find that equation (5.11) tells us that $\xi(t)$ and $\eta(t)$ are constants. Thus from equations (5.9), the unforced singular extremals are simply linear translations, which are also geodesics of the system, an example of which is shown in the first picture in Figure 9. Knowing the behaviour of both non-translated and unforced singular extremals gives us an indication of how more general singular trajectories, such as the one shown in the second picture in Figure 9, behave.

We may also note that time-optimal singular extremals are also ***abnormal*** singular extremals for an arbitrary cost function (***abnormal*** extremals are those which satisfy the maximum principle, Theorem 3.4, when we take the abnormal multiplier $p_0 = 0$). This observation may open up many other new questions regarding properties of singular extremals, and thus give one a better understanding of these interesting trajectories.

References

Bullo, F. and Lewis, A. D. [2005] Supplementary Chapters, in *Geometric Control of Mechanical Systems, Modeling, Analysis, and Design for Simple Mechanical Systems*, 49 Texts in Applied Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-22195-3, URL: `http://motion.mee.ucsb.edu/book-gcms/supplements.html`.

Ewing, G. M. [1969] *Calculus of Variations with Applications*, W. W. Norton & Co. Inc.: New York, NY, Reprint: [Ewing 1985].

— [1985] *Calculus of Variations with Applications*, Dover Publications, Inc.: New York, NY, ISBN: 978-0-486-64856-9, Original: [Ewing 1969].

Kobayashi, S. and Nomizu, K. [1963a] *Foundations of Differential Geometry*, volume 1, number 15 in Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers: New York, NY, Reprint: [Kobayashi and Nomizu 1996a].

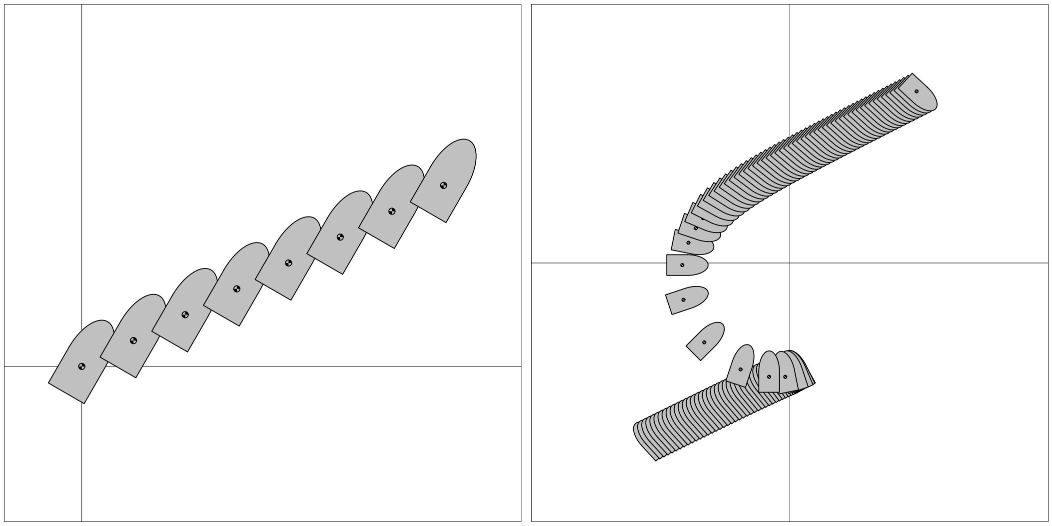

Figure 9: On the left is an unforced singular extremal given by $(\xi(t), \eta(t), \theta(t)) = (\frac{2}{25}t, \frac{1}{25}t, \frac{\pi}{3})$. On the right is the singular extremal produced by superimposing the singular trajectory of Figure 8 with the linear motion on the left.

— [1963b] *Foundations of Differential Geometry*, volume 2, number 16 in Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers: New York, NY, Reprint: [Kobayashi and Nomizu 1996b].

— [1996a] *Foundations of Differential Geometry*, volume 1, Wiley Classics Library, John Wiley and Sons: NewYork, NY, ISBN: 978-0-471-15733-5, Original: [Kobayashi and Nomizu 1963a].

— [1996b] *Foundations of Differential Geometry*, volume 2, Wiley Classics Library, John Wiley and Sons: NewYork, NY, ISBN: 978-0-471-15732-8, Original: [Kobayashi and Nomizu 1963b].

Lewis, A. D. and Murray, R. M. [1999] *Configuration controllability of simple mechanical control systems*, SIAM Review, **41**(3), Invited SIAG Review article, pages 555–574, ISSN: 0036-1445, DOI: 10.1137/S0036144599351065.

Pontryagin, L. S., Boltyanskiĭ, V. G., Gamkrelidze, R. V., and Mishchenko, E. F. [1961] *Matematicheskaya teoriya optimal′ nykh protsessov*, Gosudarstvennoe izdatelstvo fizikomatematicheskoi literatury: Moscow, Translation: [Pontryagin, Boltyanskiĭ, Gamkrelidze, and Mishchenko 1986].

— [1986] *The Mathematical Theory of Optimal Processes*, translated by K. N. Trirogoff, Classics of Soviet Mathematics, Gordon & Breach Science Publishers: New York, NY, ISBN: 978-2-88124-077-5, Original: [Pontryagin, Boltyanskiĭ, Gamkrelidze, and Mishchenko 1961].

Sussmann, H. J. [1998] *An introduction to the coordinate-free maximum principle*, in *Geometry of Feedback and Optimal Control*, edited by B. Jakubczyk and W. Respondek, Monographs and Textbooks in Pure and Applied Mathematics, pages 463–557, Dekker Marcel Dekker: New York, NY, ISBN: 978-0-8247-9068-4.

Sussmann, H. J. and Willems, J. C. [1997] *300 years of optimal control: From the brachystochrone to the maximum principle*, IEEE Control Systems Magazine, **17**(3), pages 32–44, ISSN: 1066-033X, DOI: 10.1109/37.588098.