A SECOND ORDER RIEMANNIAN VARIATIONAL PROBLEM FROM A HAMILTONIAN PERSPECTIVE

P. Crouch*
Center for Systems Science and Engineering
Arizona State University
Tempe, AZ 85287–USA
peter.crouch@asu.edu

F. Silva Leite†
Departamento de Matemática
Universidade de Coimbra
3000 Coimbra–Portugal
fleite@mat.uc.pt

M. Camarinha‡
Departamento de Matemática
Universidade de Coimbra
3000 Coimbra–Portugal
mmlsc@mat.uc.pt

Abstract

We present a Hamiltonian formulation of a second order variational problem on a differentiable manifold Q, endowed with a Riemannian metric $< \cdot, \cdot >$ and explore the possibility of writing down the extremal solutions of that problem as a flow in the space $TQ \oplus T^*Q \oplus T^*Q$. For that we utilize the connection ∇ on Q, corresponding to the metric $< \cdot, \cdot >$. In general the results depend upon a choice of frame for TQ, but for the special situation when Q is a Lie group G with Lie algebra $\mathcal{G}$, our results are global and the flow reduces to a flow on $G \times \mathcal{G} \times \mathcal{G}^* \times \mathcal{G}^*$.

Keywords: Riemannian manifolds, Lie groups, Hamiltonian equations, optimal control, variational problems.

1 Introduction

Modelling complex mechanical systems is often accomplished using the notational convenience of differential geometry and in particular Riemannian geometry and symplectic geometry. For systems without dissipation one can choose a Lagrangian approach or a Hamiltonian approach. The basic phase space is usually taken as TQ

*Work supported in part by NATO grant 4/C/94/PO.
†Work supported in part by ISR, project Praxis/2/2.1/MAT/19/94 and NATO grant 4/C/94/PO.
‡Work supported in part by ISR and project Praxis/2/2.1/MAT/19/94.

in the Lagrangian approach, and T^*Q in the Hamiltonian approach, where Q is the configuration space. Indeed, in the Hamiltonian approach, the flow is specified by a Hamiltonian vector field on T^*Q, with Hamiltonian H, and using the canonical symplectic form Ω on T^*Q. In the Lagrangian case, the flow is Hamiltonian on TQ, but in this case, one must use a suitable symplectic structure on TQ, which is obtained by pulling back Ω to TQ via a suitable bundle mapping $\Sigma : TQ \to T^*Q$. For a Lagrangian defined by kinetic plus potential energies and specified by a Riemannian metric, $< \cdot, \cdot >$ on Q, the map Σ is just the linear mapping associated to the metric and defined by

$$(\Sigma X)(Y) = < X, Y >, \quad X, Y \in \Gamma(TQ),$$

where $\Gamma(TQ)$ denotes the set of smooth vector fields on Q.

However, there is another formulation available to us in the Lagrangian case above in which one uses the Levi-Civita connection ∇ on Q, which is compatible with the Riemannian metric $< \cdot, \cdot >$, to write the system in terms of a higher order differential equation on Q. Indeed, if $L = \frac{1}{2} < \dot{q}, \dot{q} > + V(q)$, $\dot{q} \in T_q Q$, is such a Lagrangian function and Σ is defined as above we write $\Sigma \dot{q} = p \in T_q^* Q$ for the momentum of the system and $\Delta_q V \in T_q^* Q$ for the gradient of the potential function V at the point $q \in Q$. Then

$$\frac{\partial L}{\partial \dot{q}} = p.$$

We also write $\frac{D}{dt}$, to denote the covariant derivative corresponding to ∇. It follows that the Euler Lagrange equation corresponding to L is just

$$\frac{Dp}{dt} = \Delta_q V, \tag{1}$$

or

$$\frac{D^2 q}{dt^2} = \Sigma^{-1} \Delta_q V. \tag{2}$$

The corresponding Hamiltonian is of course

$$\begin{aligned} H(q,p) &= p(\dot{q}) - L(\dot{q}, q) \,|_{p=\Sigma \dot{q}} \\ &= < \dot{q}, \dot{q} > - L(\dot{q}, q) \,|_{p=\Sigma \dot{q}} \\ &= \frac{1}{2} < \dot{q}, \dot{q} > - V(q) \,|_{p=\Sigma \dot{q}} \end{aligned}$$

or

$$H(q,p) = \frac{1}{2} p(\Sigma^{-1} p) - V(q). \tag{3}$$

Now the Hamiltonian equations for H, corresponding to the Euler-Lagrange equations (1) or (2), are of course given by a vector field on T^*Q, that is a system of equations in TT^*Q. Writing down these equations depends upon fixing a system of coordinates in which to express Ω, etc. whereas the same effect has been accomplished in the Lagrangian case, reducing a system of equations in TTQ to a system of equations, either in TQ or T^*Q, through the connection ∇.

In studying the control problems for such mechanical systems, and in particular optimal control problems, one encounters even higher order bundles. For example, the following optimal control problem is a tipical situation.

$$\min_u \int_0^T \frac{1}{2} < u, u > dt \tag{4}$$

subject to

$$\dot{x} = V, \quad \frac{DV}{dt} = u \tag{5}$$

and

$$x(0) = x_0, \quad \dot{x}(0) = V_0, \quad x(T) = x_T, \quad \dot{x}(T) = V_T, \tag{6}$$

where x_0 and x_T are given points in Q, V_0 and V_t are tangent vectors to Q, at x_0 and x_T respectively.

As observed above, the system of equations (5) may already be viewed as the reduction to TQ, of a system which is viewed in the Hamiltonian setting as one in TT^*Q. To solve the optimal control problem, however, the maximum principle instructs us that extremal solutions are projections of a Hamiltonian flow in TT^*TQ. This situation is clearly very cumbersome, utilizing the canonical symplectic form on T^*TQ. In this paper we explore the possibility of writing down the extremal solutions of the problem (4) - (5) - (6) as a flow in the space $E = TQ \oplus T^*Q \oplus T^*Q$. It is clear that the dimension count of the two spaces E and TT^*TQ are the same. The idea is to utilize the connection ∇ as before, but unlike the case of flow (1) our result is dependent upon a choice of frame for TQ. Thus we obtain global results, only in the case that Q is parallelizable. For the particular case when $Q = G$, a Lie group, our results are global and the flow reduces to a flow on $G \times \mathcal{G} \times \mathcal{G}^* \times \mathcal{G}^*$ where $\mathcal{G}$ is the Lie algebra of G. We note that if Q is a Riemannian manifold then TQ has many Riemannian structures, but one, the Sasaki metric $< \cdot, \cdot >_S$, is particularly useful. Indeed, as discussed in Silva Leite, Camarinha and Crouch [14], the extremals of the problem (4) - (5) - (6) may be viewed as a sub-Riemannian geodesic flow in TQ, endowed with such a metric, but not the actual geodesic flow. Corresponding to $< \cdot, \cdot >_S$, there is a Riemannian connection ∇_S on TQ. We conjecture that we may use ∇_S to reduce the extremal Hamiltonian flow of (4) - (5) - (6) in TT^*TQ, to a flow in TTQ, without the need for a choice of frame for TQ, and retain a reasonably workable form for the equations.

Our solution to the above problem will be sought by treating it as a constrained variational problem, and utilizing the Lagrange multipliers as co-states, in a typical fashion. Indeed, extremals for the problem (4) - (5) - (6), are characterized as projections of the flow resulting from the following equation.

$$\frac{D^4 x}{dt^4} + R(\frac{D^2 x}{dt}, \frac{Dx}{dt})\frac{Dx}{dt} = 0. \tag{7}$$

(See Crouch and Silva Leite [4], [5] and Noakes, Heinzinger and Paden [10]).

This result is obtained by treating the problem as the unconstrained variational problem of minimizing the following functional subject to (6)

$$\int_0^T < \frac{D^2x}{dt^2}, \frac{D^2x}{dt^2} > dt. \tag{8}$$

Our approach follows this one with some minor modifications.

2 Preliminaries

We present here a brief review to the study of tensors on a Riemannian manifold $(Q, < \cdot, \cdot >)$. An important point being that, analogously to vector fiels, tensors can be differentiated covariantly. For more details see, for instance, Manfredo do Carmo [8] and Spivak [15]. Recall that $\Gamma(TQ)$ denotes the set of smooth vector fields on Q and $\mathcal{C}^\infty(Q)$ is the ring of real-valued smooth functions defined on Q.

A *tensor of order* k on the Riemannian manifold Q is a multilinear mapping

$$\eta : \underbrace{\Gamma(TQ) \times \cdots \times \Gamma(TQ)}_{k \text{ factors}} \longrightarrow \mathcal{C}^\infty(Q).$$

A k-*form* on Q is an alternating multilinear mapping which assigns to each $q \in Q$ an element of $(\underbrace{T_qQ \times \cdots \times T_qQ}_{k \text{ factors}})^*$. Alternating k-tensors may be identified with k-forms in a natural way through

$$\eta(X_1, \cdots, X_k)(q) = \eta(q)(X_1(q), \cdots, X_k(q)), \qquad q \in Q,\ X_i \in \Gamma(TQ).$$

In particular, tensors of order 1 are identified with one-forms (or covector fields).

A tensor of order k can be differentiated to obtain a new tensor of order $k+1$.

In what follows, η be a tensor of order 1. The *covariant differential* of η, $\nabla\eta$, is a tensor of order 2 defined by

$$(\nabla\eta)(X, Y) = Y(\eta(X)) - \eta(\nabla_Y X), \qquad \forall X, Y \in \Gamma(TQ). \tag{9}$$

For each smooth vector field Z on Q, the *covariant derivative* of η with respect to Z, denoted by $\nabla_Z\eta$, is a tensor of order 1 given by:

$$(\nabla_Z\eta)(X) = (\nabla\eta)(X, Z), \qquad \forall X \in \Gamma(TQ). \tag{10}$$

The *exterior derivative* of η, $d\eta$, is a tensor of order 2 defined by

$$d\eta(X, Y) = X(\eta(Y)) - Y(\eta(X)) - \eta([X, Y]), \quad \forall X, Y \in \Gamma(TQ). \tag{11}$$

If η is a function (i.e. a tensor of order 0), $d\eta$ is just the usual differential defined by

$$d\eta(X) = X(\eta).$$

It turns out that if the connection ∇ on Q is symmetric, that is $\nabla_X Y - \nabla_Y X = [X,Y]$, $\forall X, Y \in \Gamma(TQ)$. So, the exterior derivative of a tensor of order 1 is given by:

$$d\eta(X,Y) = (\nabla_X \eta)(Y) - (\nabla_Y \eta)(X), \quad \forall X, Y \in \Gamma(TQ). \tag{12}$$

Since we have equipped Q with the Levi-Civita connection which is symmetric, the formula (12) is always true and will be used systematically throughout the whole paper. Due to the identification of tensors and forms, all the definitions above have a natural counterpart for forms.

Now, let $\{X_1, \cdots, X_n\}$ be a frame of vector fields on Q and $\{\omega_1, \cdots, \omega_n\}$ a co-frame of covector fiels such that $\omega_k(X_j) = \delta_{kl}$. This selection must be local, unless Q is parallelizable. In terms of these frames we may write any vector field Y and covector field η along a curve $t \to x(t)$ in the following way

$$\begin{aligned} Y(x(t)) = Y(t) &= \sum_{i=1}^{n} y_i(t) X_i(x(t)) \in T_{x(t)}Q, \\ \eta(x(t)) = \eta(t) &= \sum_{i=1}^{n} \eta_i(t) w_i(x(t)) \in T^*_{x(t)}Q. \end{aligned}$$

Thus, although the X_i's and the w_i's are defined on some open set in Q, $Y(t)$ and $\eta(t)$ are only defined along the curve $x(t)$. Setting

$$\dot{x}(t) = \sum_{i=1}^{n} v_i(t) X_i(x(t)) = V(t) \in T_{x(t)}Q,$$

it follows that the covariant derivatives of Y and η, along the curve $t \to x(t)$ with velocity vector field V, are given by:

$$\begin{aligned} \frac{DY}{dt} &= \sum_{i=1}^{n} \dot{y}_i(t) X_i(x(t)) + \sum_{i=1}^{n} y_i(t)(\nabla_V X_i)(x(t)) \\ \frac{D\eta}{dt} &= \sum_{i=1}^{n} \dot{\eta}_i(t) w_i(x(t)) + \sum_{i=1}^{n} \eta_i(t)(\nabla_V w_i)(x(t)). \end{aligned}$$

We denote these expressions by the contracted forms

$$\frac{DY}{dt} = \dot{Y} + \nabla_V Y; \qquad \frac{D\eta}{dt} = \dot{\eta} + \nabla_V \eta. \tag{13}$$

It should be noted that $\dot{Y}$, $\dot{\eta}$, $\nabla_V Y$ and $\nabla_V \eta$ are all dependent on the choice of frame and co-frame, and are not invariantly defined.

3 A Variational Problem

We consider solving the optimal control problem (4) - (5) - (6) through the following variational problem:

Minimize

$$J(x,V,p_1,p_2,u)=\int_0^T(p_1(\dot{x}-V)+p_2(\frac{DV}{dt}-u)+\frac{1}{2}<u,u>)dt, \quad (14)$$

subject to the boundary conditions (6) *and the dynamics* (5).

Here $p_1(t)$, $p_2(t)$ belong to $T^*_{x(t)}Q$ and $u(t)$, $V(t)$ belong to $T_{x(t)}Q$.

We need the notion of variations of a curve and fields along curves. Let $(t,\epsilon)\rightarrow x(t,\epsilon)$, $t\in[0,T]$ and $\epsilon\in(-\rho,\rho),\rho>0$, be a parametrized family of curves satisfying

$$\begin{aligned} x(0,\epsilon)=x_0, &\quad x(T,\epsilon)=x_T, \\ \dot{x}(0,\epsilon)=V_0, &\quad \dot{x}(T,\epsilon)=V_T \end{aligned} \quad (15)$$

and $W(t)=\delta x(t)=\frac{\partial x}{\partial\epsilon}(t,0)\in T_{x(t)}Q$, $t\in[0,T]$. Thus,

$$W(0)=W(T)=0. \quad (16)$$

We similarly define variations in u,V,p_1 and p_2 as curves $(t,\epsilon)\rightarrow u(t,\epsilon)$, $V(t,\epsilon)$, $p_1(t,\epsilon)$, and $p_2(t,\epsilon)$ respectively. We may write

$$\begin{aligned} u(t,\epsilon) &= \sum_i u_i(t,\epsilon)X_i(x(t,\epsilon)) \in T_{x(t,\epsilon)}Q \\ V(t,\epsilon) &= \sum_i v_i(t,\epsilon)X_i(x(t,\epsilon)) \in T_{x(t,\epsilon)}Q \\ p_1(t,\epsilon) &= \sum_i p_i^1(t,\epsilon)w_i(x(t,\epsilon)) \in T^*_{x(t,\epsilon)}Q \\ p_2(t,\epsilon) &= \sum_i p_i^2(t,\epsilon)w_i(x(t,\epsilon)) \in T^*_{x(t,\epsilon)}Q. \end{aligned}$$

According to the notation introduced in (13) we may write

$$\frac{D}{\partial\epsilon}u(t,\epsilon)|_{\epsilon=0}=\delta u(t)+(\nabla_W u)(x(t))\in T_{x(t)}Q,$$

where

$$\begin{aligned} \delta u(t) &= \sum_i \frac{\partial u_i}{\partial\epsilon}(t,0)X_i(x(t)), \\ (\nabla_W u)(x(t)) &= \sum_i u_i(t)(\nabla_W X_i)(x(t)) \end{aligned}$$

and similarly for $\frac{DV}{\partial\epsilon}(t,\epsilon)_{\epsilon=0}$, $\frac{Dp_1}{\partial\epsilon}(t,\epsilon)_{\epsilon=0}$ and $\frac{Dp_2}{\partial\epsilon}(t,\epsilon)_{\epsilon=0}$. Thus, taking variations of the functional J in (14) we obtain

$$\delta J|_{\epsilon=0}=\int_0^T\left(p_1(\frac{D\dot{x}}{\partial\epsilon}-\frac{DV}{\partial\epsilon})+p_2(\frac{D}{\partial\epsilon}\frac{DV}{\partial t}-\frac{Du}{\partial\epsilon})+<u,\frac{Du}{\partial\epsilon}>\right)dt|_{\epsilon=0}. \quad (17)$$

Now,

$$\frac{D\dot{x}}{\partial \epsilon}|_{\epsilon=0} = \frac{DW}{\partial t}, \qquad \frac{DV}{\partial \epsilon}|_{\epsilon=0} = \delta V + \nabla_W V,$$

$$\frac{D}{\partial \epsilon}\frac{DV}{\partial t}|_{\epsilon=0} = \delta \frac{DV}{\partial t} + \nabla_W \frac{DV}{\partial t}, \qquad \frac{Du}{\partial \epsilon}|_{\epsilon=0} = \delta u + \nabla_W u.$$

At this point we need the following result.

Lemma 3.1

$$\int_0^T p_2(\delta \frac{DV}{dt})dt = \int_0^T \left(-\frac{Dp_2}{dt}(\delta V) + p_2(\nabla_{\delta V} V) \right) dt.$$

Proof

$$\begin{aligned}
\int_0^T p_2(\delta \frac{DV}{dt})dt &= \int_0^T p_2(\delta \dot{V} + \delta \nabla_V V)dt \\
&= \int_0^T p_2(\frac{D}{dt}\delta V - \nabla_V \delta V + \nabla_{\delta V} V + \nabla_V \delta V)dt \\
&= p_2(\delta V)|_0^T + \int_0^T \left(-\frac{Dp_2}{dt}(\delta V) + p_2(\nabla_{\delta V} V) \right) dt.
\end{aligned}$$

But by (15) we have $\delta V(0) = \delta V(T) = 0$ and so, the result is proved.

$\square$

Using lemma 3.1 we may rewrite (17) as

$$\begin{aligned}
\delta J|_{\epsilon=0} & \\
= & \int_0^T (p_1(\frac{DW}{dt} - \nabla_W V) + p_2(\nabla_W \frac{DV}{dt} - \nabla_W u) + < u, \nabla_W u >)dt \\
& + \int_0^T (-p_1(\delta V) - \frac{Dp_2}{dt}(\delta V) + p_2(\nabla_{\delta V} V))dt \qquad (18) \\
& + \int_0^T (-p_2(\delta u) + < u, \delta u >)dt.
\end{aligned}$$

In order for $\delta J|_{\epsilon=0}$ to vanish for all $\delta u, \delta V$ and W, the last term gives us $\Sigma u = p_2$, while the second term gives us

$$\frac{Dp_2}{dt}(X) = -p_1(X) + p_2(\nabla_X V), \quad X \in \Gamma(TM).$$

Noting the boundary condition (16) we may write $\int_0^T (p_1(\frac{DW}{dt})dt = -\int_0^T \frac{Dp_1}{dt}(W)dt$ and, according to the definition of Σ, $< u, \nabla_W u >= (\Sigma u)(\nabla_W u) = p_2(\nabla_W u)$. So, the first term in (18) yields

$$\frac{Dp_1}{dt}(X) = -p_1(\nabla_X V) + p_2(\nabla_X \frac{DV}{dt}), \quad X \in \Gamma(TM).$$

This gives our main result.

Theorem 3.2 *The extremals of the optimal control problem (4) - (5) - (6) may be expressed as solutions of the following system of equations, relative to the local choice of frame and co-frame for TQ and T^*Q:*

$$\left\{\begin{array}{lcl} \dot{x} & = & V \\ \\ \dfrac{DV}{dt} & = & \Sigma^{-1}p_2 \\ \\ \dfrac{Dp_1}{dt} & = & -p_1(\nabla V) + p_2(\nabla(\Sigma^{-1}p_2)) \\ \\ \dfrac{Dp_2}{dt} & = & -p_1 + p_2(\nabla V). \end{array}\right. \tag{19}$$

Here $\eta(\nabla Y)$, where η is a tensor and $Y \in \Gamma(TQ)$, is the function that assigns to each $X \in \Gamma(TQ)$ the real number $\eta(\nabla_X Y)$.

For future applications to Lie groups it is instructive to write out the equations in the explicit form given in the next lemma.

Lemma 3.3 *The extremal equations (19) have the following form, without the assumption 1,*

$$\left\{\begin{array}{lcl} \dot{x} & = & V \\ \\ \dot{V} & = & \Sigma^{-1}p_2 - \nabla_V V \\ \\ \dot{p}_1 & = & -dp_1(V,.) + p_2(\nabla\Sigma^{-1}p_2) - d(p_1(V)) \\ \\ \dot{p}_2 & = & -p_1 + dp_2(V,.) - 2\nabla_V p_2 + d(p_2(V)) \end{array}\right. . \tag{20}$$

Proof - From the equation (19) we have

$$\left\{\begin{array}{lcl} \dot{x} & = & V \\ \\ \dot{V} & = & \Sigma^{-1}p_2 - \nabla_V V \\ \\ \dot{p}_1 & = & -p_1(\nabla_V) - \nabla_V p_1 + p_2(\nabla\Sigma^{-1}p_2) \\ \\ \dot{p}_2 & = & -p_1 + p_2(\nabla V) - \nabla_V p_2. \end{array}\right. \tag{21}$$

Now, for $X, V \in \Gamma(TQ)$ and $\eta \in \Gamma(T^*Q)$, it follows from (12), (10) and (9) That

$$\begin{array}{rcl} d\eta(V,X) & = & (\nabla_V\eta)(X) - (\nabla_X\eta)(V) \\ \\ & = & (\nabla_V\eta)(X) - X(\eta(V)) + \eta(\nabla_X V) \\ \\ & = & (\nabla_V\eta)(X) - d(\eta(V))(X) + \eta(\nabla_X V) \end{array}$$

so,

$$d\eta(V,.) = (\nabla_V \eta) + \eta(\nabla V) - d(\eta(V)).$$

Setting $\eta = p_1$ and replacing in the expression for $\dot{p}_1$ in (21) we obtain the equation for $\dot{p}_1$ in (20).

Also from (9), (12) and the definition of exterior derivative of a function we may write

$$\begin{aligned} p_2(\nabla_X V) - (\nabla_V p_2)(X) &= d(p_2(V))(X) - (\nabla_X p_2)(V) - (\nabla_V p_2)(X) \\ &= d(p_2(V))(X) - 2(\nabla_V p_2)(X), + dp_2(V,X), \end{aligned}$$

that is, $p_2(\nabla V) - \nabla_V p_2 = d(p_2(V)) - 2\nabla_V p_2 + dp_2(V,.)$, from which the equation for $\dot{p}_2$ follows.

$\square$

The optimal control u^* is given by $u^* = \Sigma^{-1} p_2$. We notice that since $V \in T_{x(t)}Q,\ \ p_1, p_2 \in T^*_{x(t)}Q,\ \ (V, p_1, p_2) \in T_x Q \oplus T^*_x Q \oplus T^*_x Q$.

4 The Hamiltonian formalism

Our next objective is to exhibit the system of equations in Hamiltonian form, for the Hamiltonian

$$H(x, V, p_1, p_2) = \frac{1}{2} < \Sigma^{-1} p_2, \Sigma^{-1} p_2 > + p_1(V) - p_2(\nabla_V V). \tag{22}$$

Up till now we have placed no assumption on the frame or coframe. However, to ensure that the system (19) is Hamiltonian, with Hamiltonian function (22), we make the following assumption.

Assumption 1 - *The co-frame $\{w_1, \ldots, w_n\}$ consists of closed one-forms.*

From this assumption it follows that

$$\begin{aligned} dp_1 &= \Sigma p_i^1(t) dw_i(x(t)) \equiv 0, \\ dp_2 &= \Sigma p_i^2(t) dw_i(x(t)) \equiv 0. \end{aligned}$$

An immediate consequence of (12) and this assumption is that $\forall X, Y \in \Gamma(TM)$ one has

$$(\nabla_X p_1)(Y) = (\nabla_Y p_1)(X) \quad \text{and} \quad (\nabla_X p_2)(Y) = (\nabla_Y p_2)(X). \tag{23}$$

Theorem 4.1 *Under assumption 1, the system of equations (19) is Hamiltonian with Hamiltonian H in (22), in the sense that it is equivalent to the following set of equations:*

$$
\begin{cases}
\dot{x} &= \dfrac{\partial H}{\partial p_1} \\[2ex]
\dot{V} &= \dfrac{\partial H}{\partial p_2} \\[2ex]
\dot{p}_2 &= -\dfrac{\partial H}{\partial V} \\[2ex]
\dot{p}_1 &= -\dfrac{\partial H}{\partial x},
\end{cases}
\tag{24}
$$

where we view $\dfrac{\partial H}{\partial p_1} \in T_x M, \quad \dfrac{\partial H}{\partial p_2} \in T_x M, \quad \dfrac{\partial H}{\partial V} \in T_x^* M, \quad \dfrac{\partial H}{\partial x} \in T_x^* M.$

Proof - From (22) it is clear that $\dfrac{\partial H}{\partial p_1} = V$ and so the first equations in (19) and (24) are equivalent as required.

From (22) again, we have

$$\frac{\partial H}{\partial p_2} = \Sigma^{-1} p_2 - \nabla_V V.$$

So,

$$\frac{DV}{dt} = \dot{V} + \nabla_V V = \Sigma^{-1} p_2$$

and, as a consequence, $\dot{V} = \Sigma^{-1} p_2 - \nabla_V V = \dfrac{\partial H}{\partial p_2}$ as required.

We now proceed to get the third equation. From (19) and (13) respectively one gets $\dfrac{Dp_2}{dt} = -p_1 + p_2(\nabla V)$ and $\dfrac{Dp_2}{dt} = \dot{p}_2 + \nabla_V p_2$. So,

$$\dot{p}_2 = -p_1 + p_2(\nabla V) - \nabla_V p_2. \tag{25}$$

On the other hand it follows from (22), (9) and (10) that

$$\frac{\partial H}{\partial V} = p_1 + (\nabla p_2)(V) + \nabla_V p_2. \tag{26}$$

Again from (9) and (10), if X is a vector field along x,

$$p_2(\nabla_X V) = X(p_2(V)) - (\nabla_X p_2)(V) = -(\nabla_X p_2)(V), \tag{27}$$

since $p_2(V)$ does not depend on x. Replacing (27) in (26) we obtain

$$\frac{\partial H}{\partial V} = p_1 - p_2(\nabla V) + \nabla_V p_2 = -\dot{p}_2$$

and the third equation holds.

Finally, to obtain the equation for $\dot{p}_1$ we write H in (22) in the following equivalent form. Here we take into consideration that $\dfrac{DV}{dt} = \dot{V} + \nabla_V V$ and from (19) that $\dfrac{DV}{dt} = \Sigma^{-1} p_2$.

$$
\begin{aligned}
H &= \frac{1}{2} p_2(\Sigma^{-1} p_2) - p_2(\nabla_V V) + p_1(V) \\
&= \frac{1}{2} p_2(\Sigma^{-1} p_2) - p_2(\frac{DV}{dt}) + p_2(\dot{V}) + p_1(V) \\
&= -\frac{1}{2} p_2(\Sigma^{-1} p_2) + p_2(\dot{V}) + p_1(V).
\end{aligned}
$$

Thus, for any vector field X along x we have

$$X(H) = -p_2\left(\nabla_X(\Sigma^{-1} p_2)\right),$$

since $p_2(\dot{V}) + p_1(V)$ does not depend on x. Now, from (19) and (13)

$$-p_1(\nabla_V) + p_2(\nabla(\Sigma^{-1} p_2)) = \dot{p}_1 + \nabla_V p_1,$$

thus

$$\dot{p}_1 = -\nabla_V p_1 - p_1(\nabla V) + p_2(\nabla(\Sigma^{-1} p_2)).$$

However, from (9) and (10)

$$
\begin{aligned}
(\nabla_V p_1)(X) + p_1(\nabla_X V) &= (\nabla_V p_1)(X) - (\nabla_X p_1)(V) + X(p_1(V)) \\
&= dp_1(V, X),
\end{aligned}
$$

which vanishes identically by assumption 1. Thus we have

$$\dot{p}_1(X) = p_2(\nabla_X(\Sigma^{-1} p_2)) = -X(H) = -dH(X).$$

□

5 Applications to Earlier Formulations

In this section we briefly apply the equations (19) to obtain a new interpretation of existing results.

Theorem 5.1 *The system of equations (19) solves the extremal flow (7).*

We start the proof with a lemma.

Lemma 5.2 *For $Z, V \in \Gamma(TM)$, $\eta \in \Gamma(T^*M)$, one has*

$$\frac{D}{dt}\eta(\nabla V)|_Z = \frac{D\eta}{dt}(\nabla_Z V) + \eta\left(\frac{D}{dt}(\nabla_Z V) - \nabla_{\frac{DZ}{dt}} V\right).$$

Proof - Just notice that

$$\frac{d}{dt}\eta(\nabla_Z V) = \frac{D}{dt}\eta(\nabla V)\mid_Z + \eta(\nabla_{\frac{DZ}{dt}} V)$$

and simultaneously

$$\frac{d}{dt}\eta(\nabla_Z V) = \frac{D\eta}{dt}(\nabla_Z V) + \eta(\frac{D}{dt}(\nabla_Z V).$$

$\square$

Proof of Theorem 5.1 From (19) $\dfrac{Dp_2}{dt} = p_2(\nabla V) - p_1$ so, from lemma 5.2,

$$\frac{D^2 p_2}{dt^2}(Z) = -\frac{Dp_1}{dt} + \frac{Dp_2}{dt}(\nabla_Z V) + p_2\left(\frac{D}{dt}(\nabla_Z V) - \nabla_{\frac{DZ}{dt}} V\right).$$

Again, substituting from (19) and using $\frac{DZ}{dt} = \dot{Z} + \nabla_V Z$ and the symmetry of ∇ we obtain

$$\frac{D^2 p_2}{dt^2}(Z)$$

$$= p_1(\nabla_Z V) - p_2(\nabla_Z \frac{DV}{dt}) + p_2(\nabla_{\nabla_Z V} V) - p_1(\nabla_Z V) + p_2\left(\frac{D}{dt}(\nabla_Z V) - \nabla_{\frac{DZ}{dt}} V\right)$$

$$= p_2(-\nabla_Z \dot{V} - \nabla_Z \nabla_V V + \nabla_{\nabla_Z V} V + \nabla_{\dot{Z}} V + \nabla_Z \dot{V} + \nabla_V \nabla_Z V - \nabla_{\dot{Z}} V - \nabla_{\nabla_V Z} V)$$

$$= p_2(\nabla_{\nabla_Z V - \nabla_V Z} V + \nabla_V \nabla_Z V - \nabla_Z \nabla_V V)$$

$$p_2(\nabla_{[Z,V]} V + \nabla_V \nabla_Z V - \nabla_Z \nabla_V V)$$

$$= p_2(R(V,Z)V).$$

Now, since from (19) $\frac{DV}{dt} = \Sigma^{-1} p_2$ it follows that $p_2(X) = (\Sigma \dfrac{DV}{dt})(X) =< \dfrac{DV}{dt}, X >$. If X is chosen to be parallel along x, that is $\frac{DX}{dt} = 0$, we have $\dfrac{Dp_2}{dt}(X) =< \dfrac{D^2 V}{dt^2}, X >$ and $\dfrac{D^2 p_2}{dt^2}(X) =< \dfrac{D^3 V}{dt^3}, X >$. As a consequence

$$< \frac{D^3 V}{dt^3}, X >= p_2(R(V,X)V) =< \frac{DV}{dt}, R(V,X)V >$$

and using the symmetries of the curvature tensor R (see Milnor [9]), we have

$$\frac{D^3 V}{dt^3} + R(\frac{DV}{dt}, V)V \equiv 0, \tag{28}$$

which is nothing other than the equation (7) .

$\square$

We also have an expression (22) for the Hamiltonian function corresponding to the Hamiltonian system (24) . In [3], Camarinha found an invariant for the flow (28) or (7), namely the function I in the next lemma. Here we show that I is indeed H given by (22).

Lemma 5.3 *The function*

$$I = \frac{1}{2} < \frac{DV}{dt}, \frac{DV}{dt} > - < \frac{D^2V}{dt^2}, V >$$

is an invariant of the flow (28) and $I = H$, where H is the Hamiltonian function (22).

Proof

That $\frac{dI}{dt} \equiv 0$ along the flow (28), follows easily from the properties of the curvature tensor R. Since $\Sigma^{-1}p_2 = \frac{DV}{dt}$, $\Sigma^{-1}\frac{Dp_2}{dt} = \frac{D^2V}{dt^2}$ and from (19) $\frac{Dp_2}{dt}(V) = -p_1(V) + p_2(\nabla_V V)$, it easily follows that

$$\begin{aligned} H &= \frac{1}{2} < \frac{DV}{dt}, \frac{DV}{dt} > - \frac{Dp_2}{dt}(V) = \frac{1}{2} < \frac{DV}{dt}, \frac{DV}{dt} > -(\Sigma\frac{D^2V}{dt^2})(V) \\ &= \frac{1}{2} < \frac{DV}{dt}, \frac{DV}{dt} > - < \frac{D^2V}{dt^2}, V > . \end{aligned}$$

□

5.1 The Lie group case

We now specialize to the case where $Q = G$, is a compact or semisimple Lie group, with Lie algebra $\mathcal{G}$. In this case Q is parallelizable and the equations (19), and indeed the equations (24), may be given a global interpretation. In this case we also have an explicit expression for the connection corresponding to the unique bi-invariant metric on G, $\nabla_X Y = \frac{1}{2}[X, Y]$, (see, for instance, Milnor [9]).

For the Lie group G we may assume $\{X_1, \dots, X_n\}$ is a basis of left-invariant vector fields and $\{w_1, \dots, w_n\}$ is a dual basis of left-invariant one-forms. It follows that the equations (20) are indeed globally defined and we may identify V, p_1, p_2, as elements of $\mathcal{G}$, $\mathcal{G}^*$ and $\mathcal{G}^*$ respectively. However, dp_1 and dp_2 will not now vanish in general so assumption 1 cannot be made and the equations will not retain the Hamiltonian form of theorem 4.1.

At this point it is important to point out that some of the formulas in section 2 have simpler expressions when Q is a Lie group G.

If $\eta \in \mathcal{G}^*$ and ad_X is the adjoint map

$$\begin{array}{rcccl} ad_X & : & \mathcal{G} & \longrightarrow & \mathcal{G} \\ & & Y & \rightarrow & [X, Y] \end{array} \quad , \quad X \in \mathcal{G},$$

then the adjoint of ad_X is given by:

$$ad^*_X \eta(Y) = -\eta \circ ad_X(Y) = -\eta([X, Y]). \tag{29}$$

If X, Y are left-invariant vector fields on G and η is a left-invariant one-form on G, then $Y(\eta(X)) = 0, \ \forall X, Y$ and, consequently,

$$d\eta(X, Y) = -\eta([X, Y]). \tag{30}$$

Also taking into account that $\nabla_Y X = \frac{1}{2}[Y, X]$ one gets

$$\nabla_X \eta = \frac{1}{2} ad^*_X \eta. \tag{31}$$

It remains an interesting problem to identify the correct symplectic structure in this case, which will be the natural symplectic form on T^*TTG.

Lemma 5.4 *In the case of a compact or semisimple Lie group G with Lie algebra $\mathcal{G}$, the extremal equations (20) may be written in the form*

$$\begin{cases} \dot{x} = (L_x)_* V \\ \dot{V} = \Sigma^{-1} p_2 \\ \dot{p}_1 = -ad^*_V p_1 \\ \dot{p}_2 = -p_1 \end{cases}, \tag{32}$$

where $V \in \mathcal{G}, p_1, p_2 \in \mathcal{G}^*$, and $\dot{x} = (L_x)_* V$, where L_x is left translation in the Lie group G and $(L_x)_*$ is the differential of L_x at the identity e_G.

Proof - As a consequence of (30) and (31), we may write the equations (13) with $V \in \mathcal{G}, p_1, p_2 \in \mathcal{G}^*$ in the form

$$\begin{cases} \dot{x} = (L_x)_* V \\ \\ \dot{V} = \Sigma^{-1} p_2 \\ \\ \dot{p}_1 = -ad^*_V p_1 + \dfrac{1}{2} ad^*_{\Sigma^{-1}p_2} p_2 \\ \\ \dot{p}_2 = -p_1. \end{cases}$$

But from (30) and the definition of Σ,

$$ad^*_{\Sigma^{-1}p_2} p_2(X) = p_2([\Sigma^{-1}p_2, X]) = < \Sigma^{-1}p_2, [\Sigma^{-1}p_2, X] >,$$

which vanishes since $< X, [Y, Z] >=< [X, Y], Z >$.

□

Corollary 5.5 *The equations (32) imply that*

$$\dddot{V} + [V, \ddot{V}] = 0 \tag{33}$$

Proof - From equations (32), $\ddot{V} = -\Sigma^{-1}p_1$ and hence $\dddot{V} = \Sigma^{-1}ad^*_V p_1$. Therefore, $\forall X \in \mathcal{G}$,

$$<\dddot{V}, X> = (\Sigma\ \dddot{V})(X) = ad^*_V p_1(X) = p_1([V, X]) - p_1$$

and

$$< [V, \ddot{V}], X> = - < \ddot{V}, [V, X]) = -(\Sigma\ddot{V})([V, X]) = p_1([V, X]),$$

from which the result follows.

□

Equations (33) , were first written down in this generality, as a specialization of the extremal flow (28), in Crouch and Silva Leite [4] but see also Noakes, Heinzinger and Paden [10]. A specific problem in optimal control of the form (4)-(5)-(6) was treated in Bloch and Crouch [2] where an analysis was made between the Hamiltonian and the Lagrangian formulation of higher order optimal control problems. We treat the example again here in a slightly different setting.

Consider the problem:

$$\min_u \int_0^T \frac{1}{2} < u, u > dt,$$

subject to

$$\begin{cases} \dot{Q} = \Omega_1 Q \\ \\ \dot{\Omega}_1 = u \end{cases} \qquad Q \in SO(n);\ u, \Omega_1 \in so(n)\ .$$

and boundary conditions

$$Q(0) = Q_0, \quad Q(T) = Q_T, \quad \dot{Q}(0) = \dot{Q}_0, \quad \dot{Q}(T) = \dot{Q}_T.$$

Here $< A, B >= trace(A^T B)$. To solve the problem we construct the Hamiltonian

$$H(u, Q, \Omega_1, p_1, p_2) =< p_2, u > + < p_1, \Omega_1 Q > -\frac{1}{2} < u, u > . \tag{34}$$

Thus the optimal control is $u^* = p_2 \in so(n)$, from which we get

$$H = \frac{1}{2} < p_2, p_2 > + < p_1, \Omega_1 Q > .$$

Using properties of the trace of a matrix we obtain

$$\begin{aligned} \dot{p}_2 &= -\frac{1}{2}(p_1 Q^T - Q p_1^T), \\ \dot{p}_1 &= -\Omega_1^T p_1. \end{aligned} \tag{35}$$

Indeed, since

$$< p_1, \Omega_1 Q >= trace(p_1^T \Omega_1 Q) = trace(Q^T \Omega_1^T p_1) =< Q, \Omega_1^T p_1 >,$$

$$\dot{p}_1 = -\frac{\partial H}{\partial Q} = -\Omega_1^T p_1.$$

Also,

$$< p_1, \Omega_1 Q >= \frac{1}{2} trace(p_1^T \Omega_1 Q) + \frac{1}{2} trace(p_1^T \Omega_1 Q).$$

But since

$$trace(p_1^T \Omega_1 Q) = -trace(\Omega_1 Q p_1^T) = - < \Omega_1, Q p_1^T >$$

and

$$trace(p_1^T \Omega_1 Q) = -trace(\Omega_1 p_1 Q^T) =< \Omega_1, p_1 Q^T >,$$

it follows that

$$< p_1, \Omega_1 Q >= \frac{1}{2} < \Omega_1, p_1 Q^T - Q p_1^T >$$

and

$$\dot{p}_2 = -\frac{\partial H}{\partial \Omega_1} = -\frac{1}{2} < \Omega_1, p_1 Q^T - Q p_1^T > .$$

We hypothesize a solution where $p_1 = \Omega_2 Q$, with $\Omega_2 \in so(n)$. If we make this assumption it follows from (35) that

$$\begin{aligned} \dot{p}_2 &= -\Omega_2, \\ \dot{\Omega}_2 &= [\Omega_1, \Omega_2] \end{aligned}$$

and so, the full extremal equations may be written as

$$\begin{cases} \dot{Q} = \Omega_1 Q \\ \dot{\Omega}_1 = p_2 \\ \dot{p}_2 = -\Omega_2 \\ \dot{\Omega}_2 = [\Omega_1, \Omega_2] \end{cases} . \tag{36}$$

The equations (36) are precisely the equations in lemma 5.4. The Hamiltonian for the extremal flow may be written as

$$\begin{aligned} H &= \frac{1}{2} < p_2, p_2 > + < p_1, \Omega_1 Q > \\ &= \frac{1}{2} < p_2, p_2 > + < \Omega_2 Q, \Omega_1 Q > \\ &= \frac{1}{2} < p_2, p_2 > + < \Omega_2, \Omega_1 > . \end{aligned}$$

We note that with this form of H, equations (36) are not Hamiltonian, but indeed the full equations

$$\begin{cases} \dot{Q} = \Omega_1 Q \\ \dot{\Omega}_1 = p_2 \\ \dot{p}_2 = -\frac{1}{2}(p_1 Q^T - Q p_1^T) \\ \dot{p}_1 = -\Omega_1^T p_1 \end{cases}$$

are Hamiltonian with respect to the Hamiltonian

$$H = \frac{1}{2} < p_2, p_2 > + \frac{1}{2} < p_1 Q^T - Q p_1^T, \Omega_1 > .$$

References

[1] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, 2nd ed., Addison-Wesley, Reading, MA, 1978.

[2] A. Bloch and P. Crouch, Reduction of Euler Lagrange problems for constrained variational problems and relation with optimal control problems. Proc. 33^{th} IEEE CDC, 2584-2590 (1994).

[3] M. Camarinha, The Geometry of Cubic Polynomials on Riemannian manifolds, *Ph. D. Thesis*, (English translation), University of Coimbra (1996).

[4] P. Crouch and F. Silva Leite, The Dynamic Interpolation Problem on Riemannian Manifolds, Lie Groups and Symmetric Spaces, *Journal of Dynamical and Control Systems*, Vol 1, No. 2 (1995), 177–202.

[5] P. Crouch and F. Silva Leite, Geometry and the Dynamic Interpolation Problem, *Proc. American Control Conference.* Boston, 26–29 July (1991).

[6] P. Dombrowski, On the Geometry of the Tangent Bundle, *J. Reine Angew. Math.*, 210 (1962), 73-88.

[7] O. Kowalski, Curvature of the Induced Riemannian Metric on the Tangent Bundle of a Riemannian Manifold, *J. Reine Angew. Math.*, 250 (1971), 124–129.

[8] Manfredo do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[9] J. Milnor, *Morse Theory*, Ann. Math. Stud., Princeton University Press, 51, 1969.

[10] L. Noakes, G. Heinzinger and B. Paden, Cubic splines on curved spaces, *IMA Journal of Mathematical Control & Information* **6**, (1989), 465-473.

[11] K. Nomizu, *Foundations of Differential Geometry*, Interscience Tracts in Pure and Applied Mathematics, John Willey & Sons, Vol I, II (1969).

[12] S. Sasaki, On the Differential Geometry of Tangent Bundles of Riemannian Manifolds, I, *Tohôko Math. Journal*, 10 (1958), 338–354.

[13] S. Sasaki, On the Differential Geometry of Tangent Bundles of Riemannian Manifolds, II, *Tohôko Math. Journal*, 14 (1962), 146–155.

[14] F. Silva Leite, M. Camarinha and P. Crouch, Elastic curves as solutions of Riemannian and sub-Riemannian control problems. Submitted

[15] M. Spivak, *Differential Geometry*, Vol. 1, 2nd edition, Publish or Perish, Inc. Berkeley 1979.