Pontryagin's Minimum Principle for simple mechanical systems on Riemannian manifolds and Lie Groups

R. V. Iyer*

Abstract

Pontryagin's Minimum principle for optimal control had earlier been extended to non-linear systems defined on manifolds by Sussmann [1], without any additional structure such as a Riemannian metric. He pointed out that the adjoint equation can be described intrinsically using a connection along the optimal trajectory. In this paper, we describe the computation of the Hamiltonian vector field in a direct manner using the symplectic two-form defined on local coordinates on T^*TM adapted to the vertical and horizontal subspaces of TTM. This symplectic two-form is non-canonical, and its meaning and computation is akin to the non-canonical symplectic two-form on the Lie Algebra of a Lie Group. The equations turn out to be the same as those obtained in our earlier work using a calculus of variations approach [2].

1 Introduction

In this paper, we consider optimal control problems for simple mechanical systems. Such systems can be described as a vector field on the tangent bundle TM of a manifold M. There exists a natural Riemannian metric on M that is compatible with the Kinetic Energy function. Using parallel translation, one can define a frame (non-uniquely) on a co-ordinate chart of TM. The problem then is to express Pontryagin's Minimum Principle (PMP) in such frame co-ordinates. The resulting equations prove to be very useful for numerical computation. In earlier work, we obtained the first order necessary conditions using a calculus of variations approach [2]. There, we also obtained an invariant for the problem, and in this paper, we demonstrate this invariant to be the Hamiltonian function.

*R.I was supported by a NRC/AFOSR SFFP during Summer 2004 and an ASEE/AFOSR SFFP during Summer 2005; Tel: 1-806-742-2580, ext 239; Fax: 1-806-742-1112; Email: ram.iyer@ttu.edu

For affine control systems on manifolds with additional structure such as affine connection control systems with affine control, Lewis [3] described the geometry of the adjoint equation using the horizontal and vertical lifts of the horizontal and vertical subspaces on T^*TM that arise from the geodesic spray or a second-order vector field on TM. Lewis then describes the Hamiltonian vector field for such systems and a particular class of cost functions, using the cotangent lift of the geodesic spray, and obtains what he terms the adjoint equation. In this work, the set in which the input variables take values at an instant of time is considered to be state-dependent. Due to this, the adjoint equation is not entirely complete as the state dependence of the input set should introduce additional terms in the equations for the co-states [4]. We do not consider the input bounds to be related to the states in this paper. As the controlled systems and cost functions in this paper are more general than those in [3], we show how our equations for the co-states can be specialized to obtain the adjoint Jacobi equation in Section 3.4. In other work, Jurdejevic [5] and Krishnaprasad [6] have considered optimal control problems for left-invariant systems on Lie Groups, with the input variables affecting the velocity vector field on the configuration space. Koon and Marsden [7] study the problem of optimal control for nonholonomic systems with symmetry while considering the derivative of the shape space variables to be the input.

We consider simple mechanical systems with forces and moments as inputs. Sussmann [1] tackled the problem of generalizing the Pontryagin's Minimum Principle to manifolds (without any affine-connection structure), by developing the co-ordinate free Minimum principle. For computational purposes, when this principle is applied to an air-vehicle problem, one employs local co-ordinates and the equations reduce to the necessary conditions for an optimal control problem in co-ordinates. Local co-ordinates might not be the best choice possible for the real-time computation of the optimal trajectory when one has an additional Lie Group structure. This is because a suitable choice of co-ordinates depends on the initial and final conditions on the optimal control problem, which makes it unsuitable for real-time computation of optimal trajectories. If the configuration space is a connected Lie Group G with Lie Algebra $\mathcal{G}$, then one can represent any point as a product of exponentials using the exponential map, $\exp : \mathcal{G} \to G$. In general, this map is not globally one-to-one or onto, but in the case of groups that semi-direct products of a connected and compact group and a connected Abelian group, it is globally one-to-one and onto (except on a set of measure zero). Such groups arise naturally in robotics and simple mechanical systems. The first order necessary conditions yielded by the PMP can be numerically solved using the Modified Simple Shooting method as demonstrated in [2] for an optimal control problem for the rigid body.

2 Mathematical Preliminaries

In this section, we discuss the notation employed and derive some basic formulae that will be used in the next section. The mathematical notions are presented in a very concise manner, and only those notions necessary for this paper are presented. A fuller picture can be seen in references such as [8, 9, 10]. One result of this section that is the utility of choosing to parameterize the horizontal and vertical bundles of T^*TM, from the point of view of transformation of co-ordinates on the intersection of charts. This transformation property is simpler than what one would have if one chose to parameterize T^*TM directly on coordinate charts. In this case, one is said to be using *coordinate frames.* Another important purpose of this section is to express the natural symplectic two-form on T^*TM using coordinates on the horizontal and vertical bundles of T^*TM.

A a manifold $\{M, \{U_\alpha; \alpha \in I\}\}$ is separable Hausdorff space together with a collection of open sets that cover it with the condition that $\emptyset$, $M \in \{U_\alpha\}$, and that is locally homeomorphic to $\mathbb{R}^n$. Here, I is an index set. So given any $q \in M$ with $q \in U$ where U is an open neighborhood of M, there is some homeomorphism $\varphi : U \rightarrow \mathbb{R}^n$. We will assume that n is a constant that does not depend on q. The collection $\{U_\alpha, \varphi_\alpha\}$ is called a set of coordinate charts for M, and they satisfy the condition that $\varphi_\alpha \circ \varphi_\beta^{-1} : U_\alpha \cap U_\beta \rightarrow U_\alpha \cap U_\beta$ is a homeomorphism. We will consider a C^∞ manifold, where the map $\varphi_\alpha \circ \varphi_\beta^{-1}$ is a C^∞ diffeomorphism.

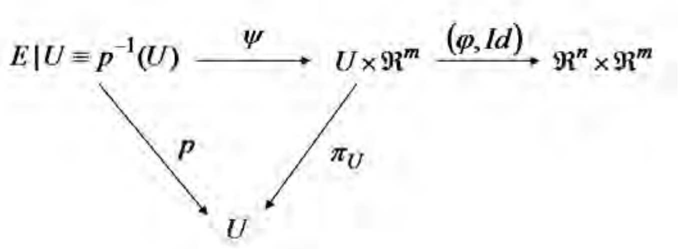

Figure 1: Local Trivialization for a Vector Bundle.

A vector bundle $\{E, M, p, \mathbb{R}^m, \psi, \{g_{\alpha\beta}; \alpha, \beta \in I\}\}$ is collection of sets E, M; a map $p : E \rightarrow M$; a *trivializing map* $\psi : p^{-1}(U) \rightarrow U \times \mathbb{R}^m$ that is fiber respecting in the sense that $\psi(p^{-1}(x)) = \{x\} \times \mathbb{R}^m$. Two vector bundle charts (U_α, ψ_α) and (U_β, ψ_β) satisfy the *compatibility condition* $(\psi_\alpha \circ \psi_\beta^{-1})(q, v) = (q, g_{\alpha\beta}(x)\, v)$. The trivializing maps $\{\psi_\alpha; \alpha \in I\}$ also have to satisfy a *co-cycle* condition [9, 8]. By the proof of the vector bundle construction theorem [8], we can say that a point on the vector bundle E is an equivalence class $[q, \alpha, v]$ where $q \in U_\alpha$ and $v \in \mathbb{R}^m$ - here, two points (q, α, v) and (r, β, w) are considered equivalent if and only if $q = r$ and $w = g_{\beta\alpha}(q)v$.

The fiber of the point $q \in M$ given by $E_p = p^{-1}(q)$ is a vector space isomorphic to $\mathbb{R}^m$ with $a\,[q,\alpha,v] + b[q,\beta,w] = [q,\alpha, a\,v + b\,g_{\alpha\beta}(q)w]$. In the case of the tangent bundle $E = TM$, the map $g_{\alpha\beta}$ is obtained from the coordinate maps φ_α, φ_β according to: $g_{\alpha\beta}(q) = D(\varphi_\alpha \circ \varphi_\beta^{-1})(\varphi_\beta(q))$, where D denotes the Frechét derivative. Coordinate chart maps for the vector bundle are obtained simply as: $\Phi_\alpha = (\varphi_\alpha, Id) \circ \psi$ as shown in Figure 1. This describes the differential manifold structure of E. For the tangent bundle, the map p is usually denoted by π_M.

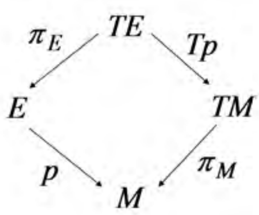

Figure 2: Tangent Bundle to a Vector Bundle has two compatible vector bundle structures.

The tangle bundle to a vector bundle described above is itself a vector bundle over TM as well as over E as shown in Figure 2. A point in TE is an equivalence class $\{[q,\alpha,v,\xi,w]\}$, where $q \in U_\alpha$, $\xi \in \mathbb{R}^n$ and v, $w \in \mathbb{R}^m$, with two points $(q,\alpha,v_\alpha,\xi_\alpha,w_\alpha)$ and $(r,\beta,v_\beta,\xi_\beta,w_\beta)$ considered equivalent if and only if

$$q = r;\quad v_\beta = g_{\beta\alpha}(q)v_\alpha;\quad \xi_\beta = h_{\beta\alpha}(q)\xi_\alpha;\quad \text{and} \quad w_\beta = (D(g_{\beta\alpha}(q))\xi_\alpha)v_\alpha + g_{\beta\alpha}(q)w_\alpha, \tag{1}$$

where $h_{\beta\alpha}(q) = D(\varphi_\beta \circ \varphi_\alpha^{-1})(\varphi_\alpha(q))$. The two vector bundle structures are compatible in the sense that $p \circ \pi_E = \pi_M \circ Tp$, where π_E and π_M are both projection maps described earlier. This implies that the fiber at a point $q \in M$ in the vector bundle $(TE, M, p \circ \pi_E, \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m)$ is identical to the fiber in $(TE, M, \pi_M \circ Tp, \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m)$ which leads to the same trivializing map and chart compatibility condition.

Given a vector bundle $(E, M, p, \mathbb{R}^m)$ one can define the *vertical bundle* $VE \subset TE$ to be null space Null(Tp). As the map $Tp : TE \to TM$ is described in a local trivialization by $Tp(q,v,\xi,w) = (q,\xi)$, a point on the vertical bundle (in the same local trivialization) is given by $(q,v,0,w)$. After checking the compatibility condition, it can be easily confirmed that $(VE, E, \pi_E, \mathbb{R}^m)$ is a vector bundle. A linear Ehresmann *connection* on the vector bundle $(E, M, p, \mathbb{R}^m)$ is a fiber-linear, vector valued one form Φ that acts on vector-fields on M and yields sections in the vertical bundle VE, such that Φ restricted to VE is the identity map and Image(Φ) $= VE$. Hence, in a local trivialization, the linear connection is described by: $\Phi(q,v,\xi,w) = (q,v,0,w + \Gamma^i_{jk}(q)\,\xi^j\,v^k\,e_i)$, where $\{e_i\}$ is the basis for $T_{\varphi(q)}\mathbb{R}^m$ and repeated indices are summed according to the Einstein convention. Due to the fact that $HE = \text{Null}(\Phi)$ has a constant rank n, we have a decomposition $TE = \text{Image}(\Phi) \oplus \text{Null}(\Phi) = VE \oplus HE$. Thus, in a local trivialization a point $(q, v, \xi, w) \in TE$ can be written as the sum of $(q, v, 0, w + \Gamma^i_{jk}(x)\, \xi^j\, v^k\, e_i) \in VE$ and $(q, v, \xi, -\Gamma^i_{jk}(x)\, \xi^j\, v^k\, e_i) \in HE$. The *freeing map* $\mathcal{F} : VE \to E \times_M E)$ given locally by $(q, v, 0, \bar{w}) \mapsto ((q, v), (q, \bar{w}))$ and the projection $\mathcal{G} = Tp\Big|_{HE} : HE \to E \times_M TM$ given locally by $(q, v, \xi, -\Gamma^i_{jk}(q)\, \xi^j\, v^k\, e_i) \mapsto ((q, v), (q, \xi))$ are natural. The *vertical lift* map $vlft : E \times_M E \to VE$ given locally by $((q, v), (q, \bar{w})) \mapsto (q, v, 0, \bar{w})$ is the inverse of the map $\mathcal{F}$ and a right-inverse of the map Tp. Similarly, one can define the horizontal lift $hlft : E \times_M TM \to HE$ given locally by $((q, v), (q, \xi)) \mapsto (q, v, \xi, -\Gamma^i_{jk}(q)\, \xi^j\, v^k\, e_i)$, that is the inverse of $\mathcal{G}$ and another right inverse of Tp.

Next, we describe the *Koszul* connection on a fiber bundle (E, M, p). Let $\aleph(M; E)$ denote the space of sections from the base manifold M to the total manifold E. If $\zeta \in \aleph(M; E)$ and $X \in \aleph(M; TM)$, then the Koszul connection is a map $\nabla : \aleph(M; E) \times_M \aleph(M; TM) \to \aleph(M; E)$ given by $\nabla_X(\zeta) := \pi_2 \mathcal{F} \circ \Phi \circ T\zeta(X)$. In words, it is directional derivative of ζ in the direction X projected to VE using the Ehresmann connection; then onto E using the freeing map, and finally onto the fiber. In a local trivialization we have: $\nabla_{(x,\xi)}(q, v) = (q, (\xi(v^i) + \Gamma^i_{jk}(q)\, \xi^j\, v^k)\, e_i)$. It is easy to check that Koszul connection satisfies the properties: (i) $\nabla_{fX}\zeta = f\, \nabla_X \zeta$ where $f \in C^\infty(M; \mathbb{R})$; (ii) $\nabla_{X_1 + X_2}\zeta = \nabla_{X_1}\zeta + \nabla_{X_2}\zeta$; (iii) $\nabla_X(h\, \zeta) = X(h)\, \zeta + h\, \nabla_X \zeta$ where $h \in C^\infty(E; \mathbb{R})$; (iv) $\nabla_X(\zeta_1 + \zeta_2) = \nabla_X \zeta_1 + \nabla_X \zeta_2$, and it is thus a *derivation*.

For a vector bundle $(E, M, p, \mathbb{R}^m)$ the dual vector bundle $(E^*, M, p_{dual}, \mathbb{R}^{m*})$ can be constructed in a straight forward way so that $E^*_x = p^{-1}_{dual}(q)$ is the dual of $E_q = p^{-1}(q)$, where $q \in M$ [8]. As in the case of E, a point on E^* is an equivalence class $[q, \alpha, \mu]$ where $q \in U_\alpha$ and $\mu \in \mathbb{R}^{m*}$. Denote the inner-product between a vector and co-vector by in the chart $(U_\alpha, \varphi_\alpha)$ by $\langle \cdot, \cdot \rangle$. As the inner-product should be the same in any chart, we need $\langle \mu_\alpha, v_\alpha \rangle = \langle \mu_\beta, v_\beta \rangle = \langle \mu_\beta, g_{\beta\alpha} v_\alpha \rangle = \langle g_{\beta\alpha} \mu_\beta, v_\alpha \rangle$ implying $\mu_\alpha = g_{\beta\alpha}\mu_\beta$. This yields the compatibility condition for the dual bundle. The above discussion works for the tangent vector bundle, the vertical bundle, and the horizontal bundle $(HE, E, \pi_E, \mathbb{R}^n)$, and we get the dual bundles $(T^*E, E, \pi_{E,dual}, \mathbb{R}^n \times \mathbb{R}^m)$ and $(V^*E, E, \pi_{E,dual}, \mathbb{R}^m)$ and the dual horizontal bundle $(H^*E, E, \pi_{E,dual}, \mathbb{R}^n)$. By the isomorphism $VE \equiv E \times_M E$ and $HE \equiv E \times_M TM$ discussed earlier, we have $V^*E \equiv E \times_M E^*$ and $H^*E \equiv E \times_M T^*M$.

Finally, we discuss *vertical one-forms* or sections of the vector bundle (V^*E, E, π) [10]. Due to the isomorphism $V^*E \equiv E \times_M E^*$, a vertical one-form $(q, v, 0, p_2)$ can be written in a local trivialization simply as $((q, v), (q, p_2)$. The isomorphism mentioned above can be seen as the result of the adjoint $\mathcal{F}^* : E \times_M E^* \to V^*E$ of the freeing map $\mathcal{F} : VE \to E \times_M E$ (see Figure 3 for the special case $E = TM$ that is discussed next).

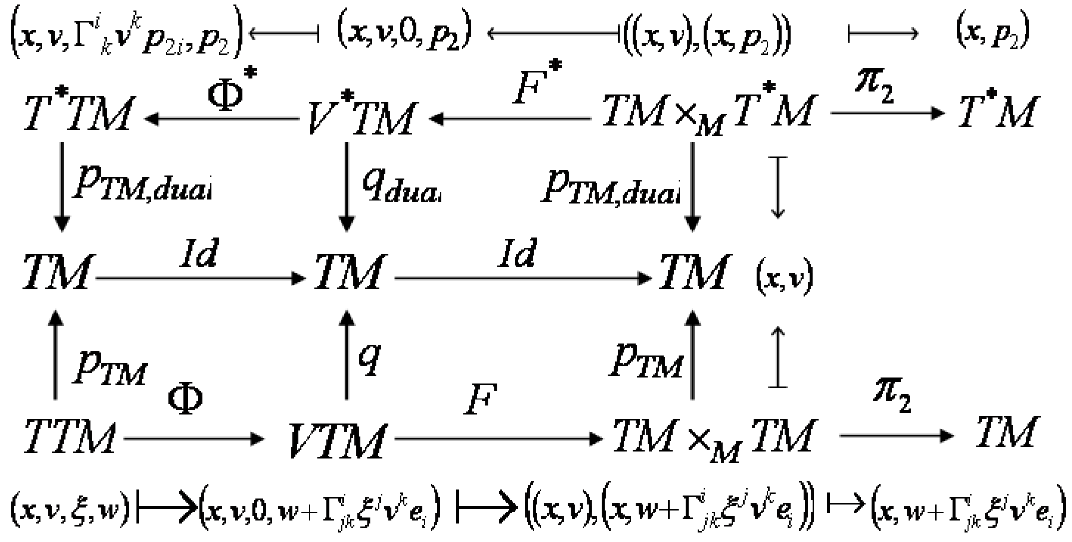

Figure 3: The Ehresmann connection map Φ, the Freeing map $\mathcal{F}$ and their adjoints.

2.1 Pull-back of the Liouville one-form to T^*TM

We now turn our attention to the special case $E = TM$ that is presented in Figure 3. A section σ of the bundle $p_{TM,dual} : TM \times_M T^*M \to TM$ is a map $\sigma : TM \to TM \times T^*M$. Suppose that in local coordinates $\sigma(q, v) = (q, p_2) = p_{2i}\, e^i$ where $\{e^i(q),\, i = 1, \cdots, n\}$ is the dual of the frame $\{e_i(q),\, i = 1, \cdots, n\}$ defined at T_qM. Then the pull-back $\bar{\sigma} \in \aleph(TM, V^*TM)$ is then given in local coordinates by:

$$\bar{\sigma}(q, v) = (0, p_2) = p_{2i}\, dv^i. \tag{2}$$

We can pull-back σ using the adjoint map $\Phi^* : V^*TM \to T^*TM$ to yield a section $\hat{\bar{\sigma}} \in \aleph(TM, T^*TM)$ that is given in coordinates by:

$$\hat{\bar{\sigma}}(q, v) = (\Gamma^i{}_k v^k p_{2i}, p_2) = \Gamma^i_{jk} v^k p_{2i} e^j + p_{2i}\, dv^i. \tag{3}$$

It is easy to check that the range of $\Phi^*(V^*TM)$ (denoted by $\mathcal{R}(\Phi^*)$ in Figure 4) is the annihilator of HTM.

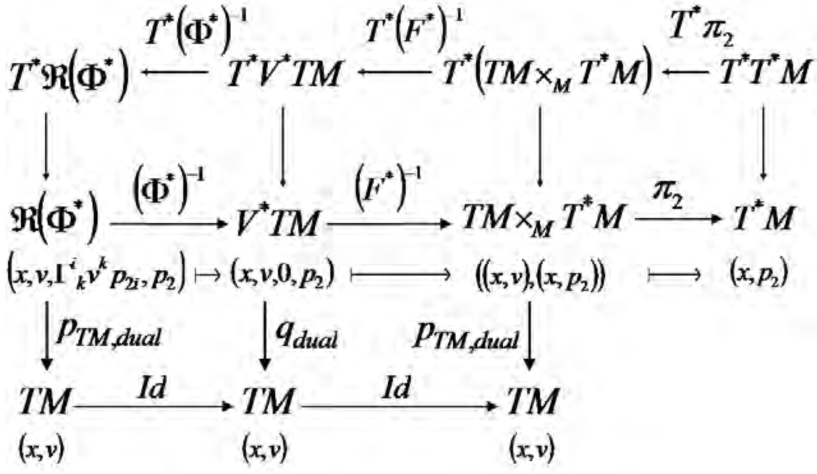

Figure 4: Pullback of the Liouville one-form to $\mathcal{R}(\Phi^*)$.

The Liouville (or canonical) one-form θ_0 on T^*M is the unique one-form that satisfies $T^*\beta \circ \theta_0 = \beta$ for any one-form β on M [11, 10]. In local coordinates on T^*M, it is given by $\theta_0 = (q, p_2, p_2, 0)$ or $\theta_0 = p_{2i}\, e^i$. The map Φ^* is also an isomorphism of its domain onto its range by its very definition. Hence: $\Phi^{*-1} : \mathcal{R}(\Phi^*) \to V^*TM$ is well-defined and we have a well-defined chain of maps: $\pi_2 \circ \mathcal{F}^{*-1} \circ \Phi^{*-1} : \mathcal{R}(\Phi^*) \to T^*M$ that we can use to pull back the Liouville one-form from $\aleph(T^*M, T^*T^*M)$ to $\aleph(T^*TM, T^*V^*TM$ where it is described in local coordinates as $p_{2i} dv^i$. This one-form can be pulled back to $\aleph(T^*TM, T^*\mathcal{R}(\Phi^*))$ via $T^*\Phi^{*-1}$ where it is given locally by:

$$\Theta_v = \Gamma^i_{jk} v^k p_{2i} e^j + p_{2i}\, dv^i. \tag{4}$$

Next, we discuss a similar construction for the horizontal bundle HTM that is shown in Figures 5 and 6. Denote $\chi = Id - \Phi$ so that $\xi : TTM \to HTM$ is the complementary projection to the horizontal bundle. Suppose that a section γ of the bundle $p_{TM,dual} : TM \times_M T^*M \to TM$ is described in local coordinates by $\gamma(q, v) = (q, p_1) = p_{1i}\, e^i$. Then the pull-back $\bar{\gamma} \in \aleph(TM, H^*TM)$ is then given in local coordinates by:

$$\bar{\gamma}(q, v) = (p_1, 0) = p_{1i}\, e^i. \tag{5}$$

We can pull-back γ using the adjoint map $\chi^* : V^*TM \to T^*TM$ to yield a section $\hat{\bar{\gamma}} \in \aleph(TM, T^*TM)$ that is given in coordinates by:

$$\hat{\bar{\gamma}}(q, v) = (p_1, 0) = p_{1i} e^i. \tag{6}$$

One can check that the range of $\chi^*(H^*TM)$ (denoted by $\mathcal{R}(\chi^*)$ in Figure 6) is the annihilator of VTM.

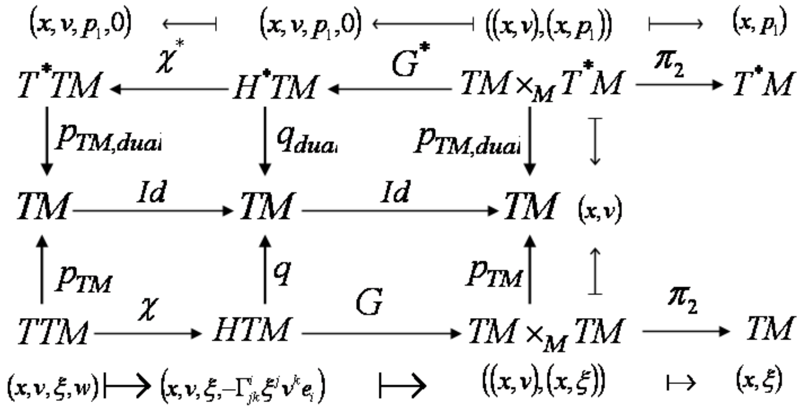

Figure 5: The projection map χ, the Freeing map $\mathcal{G}$ and their adjoints.

Again, in local coordinates on T^*M, the Liouville one-form is given by $\theta_0 = (q, p_1, p_1, 0)$ or $\theta_0 = p_{1i}\, e^i$. Similar to Φ^*, the map χ^* is an isomorphism of its domain onto its range by its very definition. Hence: $\chi^{*-1} : \mathcal{R}(\chi^*) \to H^*TM$ is well-defined and we have a well-defined chain of maps: $\pi_2 \circ \mathcal{G}^{*-1} \circ \chi^{*-1} : \mathcal{R}(\chi^*) \to T^*M$ that we can use to pull back the Liouville one-form from $\aleph(T^*M, T^*T^*M)$ to $\aleph(T^*TM, T^*H^*TM)$ where it is described in local coordinates as $p_{1i}e^i$. This one-form can be further pulled back to $\aleph(T^*TM, T^*\mathcal{R}(\chi^*))$ via $T^*\chi^{*-1}$ where it is given locally by:

$$\Theta_h = p_{1i}\, e^i. \tag{7}$$

In some coordinate chart, let $\bar{p} = (\bar{p}_1, \bar{p}_2) \in T^*_{(q,v)}TM$. We can express $\bar{p}$ as:

$$\bar{p} = (p_1, 0) + (\Gamma^i_k v^k p_{2i}, p_2) \quad \text{where} \;\; p_1 = \bar{p}_1 - \Gamma^i_k v^k \bar{p}_{2i}, \;\; \text{and} \;\; p_2 = \bar{p}_2. \tag{8}$$

Clearly the transform $(\bar{p}_1, \bar{p}_2) \mapsto (p_1, p_2)$ is one-one and onto, with inverse $\bar{p}_1 = p_1 + \Gamma^i_k v^k p_{2i}$, $\bar{p}_2 = p_2$. The coordinates (p_1, p_2) have much nicer transformation properties at the intersection of two charts. This can be seen as follows. Consider a point $x \in U_\alpha \cap U_\beta$ and two equivalent points $(q, \alpha, v_\alpha, \xi_\alpha, w_\alpha)$, $(x, \beta, v_\beta, \xi_\beta, w_\beta) \in TTM$. These points are related according to (1). Now, consider the "vertical" vectors $\Phi(q, \alpha, v_\alpha, \xi_\alpha, w_\alpha) = (q, \alpha, v_\alpha, 0, w_\alpha + \Gamma_{\alpha,jk}\xi^j_\alpha v^k_\alpha)$ and $\Phi(q, \beta, v_\beta, \xi_\beta, w_\beta) = (q, \beta, v_\beta, 0, w_\beta + \Gamma_{\beta,jk}\xi^j_\beta v^k_\beta)$ that must also be equivalent. By (1) we have:

$$w_\beta + \Gamma_{\beta_{jk}} \xi^j_\beta v^k_\beta = g_{\beta\alpha}(w_\alpha + \Gamma_{\alpha_{jk}} \xi^j_\alpha v^k_\alpha). \tag{9}$$

Now consider any one coordinate chart. The inner product between the co-vector $(\bar{p}_1, \bar{p}_2) \in$

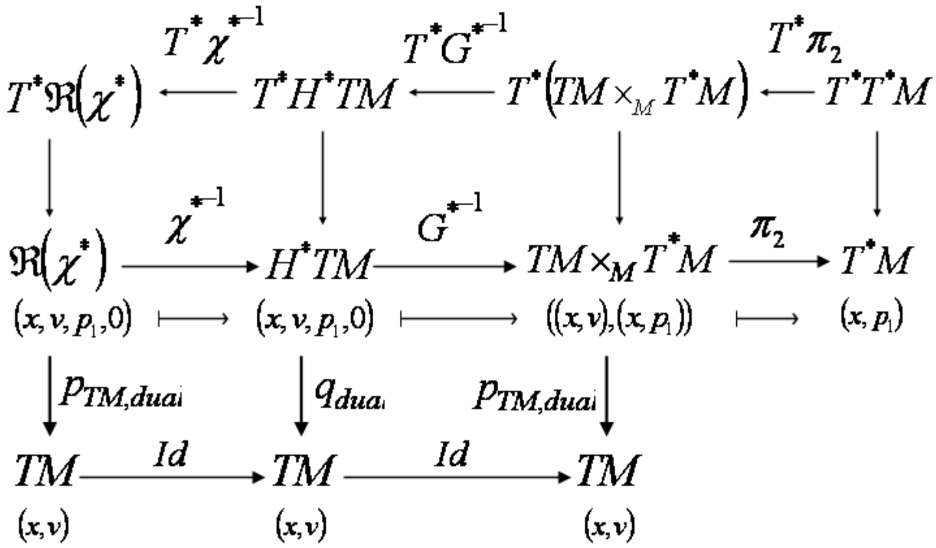

Figure 6: Pullback of the Liouville one-form to $\mathcal{R}(\xi^*)$.

$T^*_{(q,v)}TM$ and the vector $(\xi, w) \in T_{(q,v)}TM$ can be seen to be:

$$\langle(\bar{p}_1, \bar{p}_2), (\xi, w)\rangle = \langle \bar{p}_1 - \Gamma^i_{\ k}\bar{p}_{2i}v^k, \xi\rangle + \langle \bar{p}_2, w + \Gamma_{jk}\xi^j v^k\rangle = \langle p_1, \xi\rangle + \langle p_2, w + \Gamma_{jk}\xi^j v^k\rangle \tag{10}$$

by direct computation. Equations (1) and (9) together with (10) imply that in the intersection $U_\alpha \cap U_\beta$, we have:

$$p_{\beta_1} = h_{\alpha\beta}p_{\alpha_1} \quad \text{and} \quad p_{\beta_2} = g_{\alpha\beta}p_{\alpha_2}, \tag{11}$$

where we explicitly mentioned the coordinate chart where the points belong.

Definition 2.1 *The map $\Theta_0 \in \aleph(T^*TM, T^*T^*TM)$ defined by:*

$$\Theta_0 \triangleq \Theta_h + \Theta_v \tag{12}$$

is the Liouville one-form on T^*TM, where $T^*\beta$ is the cotangent lift of β.

The reason for calling Θ_0 as the Liouville one-form is the lemma below.

Lemma 2.1 *Θ_0 is the unique one-form on T^*TM such that $T^*\beta \circ \Theta_0 = \beta$ for any local one-form $\beta \in \aleph(TM, T^*TM)$.*

Proof: The lemma could be proved by using Proposition 6.3.2 on page 152 of [12] by noting that Θ_h and Θ_v are obtained through cotangent lifts. But, we give a different proof that yields some more insight. In some coordinate chart, let $(\bar{p}_1, \bar{p}_2) \in T^*_{(q,v)}TM$. Then:

$$\begin{aligned}\Theta(q, v, \bar{p}_1, \bar{p}_2) &= \Theta_h(q, v, p_1, 0) + \Theta_v(q, v, \Gamma^i_k v^k p_{2i}, p_2) \quad \text{(by definition)} \\ &= p_{1j} e^j + \Gamma^i_k v^k p_{2i} + p_{2i} dv^i \quad \text{(by (4) and (7))} \\ &= \bar{p}_{1i} e^i + \bar{p}_{2i} dv^i \quad \text{(by (8))}.\end{aligned}$$

The last equality shows that the action of $\Theta(q, v, \bar{p}_1, \bar{p}_2)$ on a vector $v \in T_{(q,v,\bar{p}_1,\bar{p}_2)}T^*TM$ can be described as:

$$\langle \Theta(q, v, \bar{p}_1, \bar{p}_2), v \rangle = \langle (\bar{p}_1, \bar{p}_2), Tp_{TM} \cdot v \rangle, \tag{13}$$

where $p_{TM} : T^*TM \to TM$ is the projection, and $Tp_{TM} : TT^*TM \to TTM$ is the tangent map of p_{TM}. The lemma then follows by Proposition 6.2.2 of [12].

□

Henceforth, we will use the coordinates (p_1, p_2) instead of $(\bar{p}_1, \bar{p}_2)$ (see (8)) due to their nicer transformation properties in coordinate charts. As Θ is the Liouville one-form, the symplectic two-form $\Sigma \in \aleph(T^*TM; \Lambda^2(T^*T^*TM))$ is given intrinsically by: $\Sigma = -d\,\Theta$ and in local coordinates:

$$\Sigma(q, v, p_1, p_2) = e^i \wedge dp_{1i} - p_{1i}\, de^i + dv^i \wedge dp_{2i} \,-\, p_{2i}\, v^k\, d\Gamma^i{}_k + v^k\, \Gamma^i{}_k \wedge dp_{2i} + p_{2i}\, \Gamma^i{}_k \wedge dv^k. \tag{14}$$

Now using the fact that $de^i = -\Gamma^i{}_k \wedge e^k$ for an orthonormal co-frame $\{e^i;\, i = 1, \cdots, n\}$ (due to the connection being torsion-free) and $d\Gamma^i{}_k = \Omega^i{}_k - \Gamma^i{}_j \wedge \Gamma^j{}_k$ where $\Omega^i{}_k$ is the curvature tensor [15], we have:

$$\Sigma(q, v, p_1, p_2) = e^i \wedge (dp_{1i} - p_{1j}\Gamma^j{}_i) - v^k p_{2i} \Omega^i{}_k + (dv^i + \Gamma^i{}_k v^k) \wedge (dp_{2i} - p_{2j}\Gamma^j{}_i). \tag{15}$$

3 Pontryagin's Minimum Principle on Riemannian Manifolds

In this section, we obtain expressions for the first order necessary conditions yielded by the PMP for a simple mechanical system on a Riemannian manifold. The general form of the PMP in a coordinate free setting was described by that is given in an abstract form in [1]. Here we specialize those results to simple mechanical systems that are *time-varying or Carathéodory second order systems.* We emphasize that the time-variation of the systems is not due a time-varying connection, but rather due to the control input which is a function of time. In applications such as trajectory design for a space shuttle or hypersonic air vehicle during ascent, where there is a change in the mass and moment of inertia of the air vehicle, one should consider the vector field with the time variation as part of the vector field with the control input. In the following subsection, we will specialize the notation and the definitions employed in Sussmann [1] to time-varying second order systems.

3.1 Notation and Definitions for the Optimal Control Problem

If S is a set (for example, $S = M$, TM, T^*TM etc where M is a smooth manifold), a *time-varying* map on S is a map whose domain is $S \times I$ where I is some interval on $\mathbb{R}$. If f is a time-varying map on S with domain $S \times I$, then I is called the time-domain of f denoted by $\mathcal{TD}(f)$.

3.1.1 Carathéodory Functions, Vector fields, Integral Curves and Controlled Curves

A *Carathéodory function* (CF) on S is a time-varying function f such that (i) $f(\cdot, t)$ is continuous for every $t \in \mathcal{TD}(f)$, and (ii) $f(x, \cdot)$ is Lebesgue measurable for every $x \in S$. The notation $\mathcal{CF}(S)$ denotes the set of all Carathéodory functions, while $\mathcal{CF}(I, S)$ denotes those functions in $\mathcal{CF}(S)$ with time-domain I. If $f \in CF(S)$, then f is called *locally integrably bounded* (LIB) if for every compact subset K of S there exists a $g \in L^1_{loc}(\mathcal{TD}(f))$ such that $|f(x, t)| \leq g(t)$ for all $(x, t) \in K \times \mathcal{TD}(f)$. In our case, S will be either $TM \times \mathcal{U}$ or $T^*TM \times \mathcal{U}$ where $\mathcal{U} \subset \mathbb{R}^m$ and M is a Riemannian manifold. Hence S has the structure of a metric space with some distance function, say, d. A function $f \in \mathcal{CF}(S)$ is called *locally Lipschitz* (LL) if every $f(\cdot, t)$ is locally Lipschitz, and *locally integrably Lipschitz* (LIL) if it is LIB and LL, and for every compact subset K of S the Lipschitz constant for f can be chosen to be in $L^1_{loc}(\mathcal{TD}(f))$. Sussmann [1] observes that definitions of a LL and LIL CF on S only depends on the class of locally equivalent metrics on S, and so they are well defined when $S = TM$ or T^*TM where M is a Riemannian manifold.

If $f \in CF(S)$ where S is a manifold, then f is said to be of class C^k if $f(\cdot, t) \in C^k(S)$ for every $t \in \mathcal{TD}(f)$. f is said to be locally integrably of class C^k (denoted LIC^k) if f is of class C^k and the function $X_1 X_2 \cdots X_k f$ is LIB for every k-tuple $(X_1, \cdots, X_k)$ of smooth vector fields on S.

A *Carathéodory vector field* (denoted CVF) on S is a time-varying map F on S such that (i) $F(x, t) \in T_x S$ fo every (x, t) in the domain of F, and (ii) $F(f) \in CF(S)$ for every $f \in C^\infty(S)$. Denote $CVF(I, S)$ to be the set of CVF's with time domain I and set $CVF(S) = \underset{I}{\cup} CVF(I, S)$. If $p : TM \to M$ denotes the projection operator, then a *second-order Carathéodory vector field* on M (denoted by $SOCVF(M)$) is a $CVF(TM)$ that satisfies $Tp \circ F = Id$. A curve on S is a continuous map $c : I \to S$, where I is an interval. A curve c on S is *locally absolutely continuous* (denoted by LAC), *locally Lipschitz* (denoted by LL), *of class* C^k, if $f \circ c$ is LAC, LL, or of class C^k for every $f \in C^\infty(S)$. An *arc* is a LAC curve such that the domain is compact. The sets $CRV(I,S)$, $CRC_{LAV}(I,S)$, $CRV(S)$, $CRV_{LAC}(S)$, $ARC(S)$ stand for the sets just described.

An *integral curve* (IC) of an $F \in CVF(S)$ is a LAC curve c such that the Domain$(c) \subset \mathcal{TD}(F)$ and $\dot{c} = F(c(t),t)$ for almost all $t \in$ Domain(c). If $F \in SOCVF(M)$, then it is clear from the above discussion that $c \in IC(F)$ if and only if in any coordinate chart for TM, c is given in local coordinates by $c(t) = (q(t), V(t))$ where $V(t) \in T_{q(t)}M$ and q is the integral curve of $V \in CVF(M)$. By the Carathéodory existence and uniqueness theorems, given a LIL CVF F on S (respectively, a SOCVF F on M); and a point $(\bar{x}, \bar{t}) \in S \times \mathcal{TD}(F)$ (respectively, a point $(\bar{q}, \bar{V}, \bar{t}) \in TM \times \mathcal{TD}(F)$) there exists an integral curve c in S (respectively, an integral curve $c = (q, V)$ in TM) such that (i) $c(\bar{t}) = \bar{x}$ (respectively, $q(\bar{t}) = \bar{q}$, $V(\bar{t}) = \bar{V}$); (ii) Domain(c) is a neighborhood of $\bar{t}$ relative to $\mathcal{TD}(F)$; (iii) for any two IC's c_1 and c_2 with (i) and (ii) satisfied, we have $c_1(t) = c_2(t)$ for $t \in$ Domain$(c_1) \cap$ Domain(c_2) [1].

A *controlled curve* in S is a pair $\gamma = (c, F)$ such that $F \in CVF(S)$ and $c \in IC(F)$. A LIL-controlled curve is one where F is LIL.

3.1.2 Hamiltonian Systems

The symplectic manifold T^*S is endowed with a canonical symplectic form $\Sigma \triangleq -d\theta_0$ where θ_0 is the Liouville one-form on T^*S. For every CVF F and CF L on S we can associate the pair (F, L) the *Hamiltonian* $H^{F,L} \in CF(T^*S)$ with time domain $\mathcal{TD}(H) = \mathcal{TD}(F) \cap \mathcal{TD}(L)$ defined by:

$$H^{F,L}(x, z, t) \triangleq \langle z, F(x,t) \rangle + L(x,t), \tag{16}$$

where $x \in S$, $z \in T^*_x S$, $t \in \mathcal{TD}(H)$. The Hamiltonian H is LIB, LIC^k, or LIL if and only if both L and F are LIB, LIC^k, or LIL. If $H \in CF_{C^k}(T^*S)$; $k \geq 1$, we can associate with H a *Hamiltonian vector field* $X_H \in \aleph^{k-1}(T^*S, TT^*S)$ with time domain $\mathcal{TD}(H)$ defined by:

$$\langle dH(x,t), w \rangle = \Sigma_x(X_H(x,t), w) \quad \text{for all } \; x \in T^*S, \; t \in \mathcal{TD}(H), \text{ and } w \in T_x T^*S. \tag{17}$$

If $F \in CVF_{LIC^k}(T^*S)$ and $L \in CF_{LIC^k}(T^*S)$, we have $X_H \in CVF_{LIC^{k-1}}(T^*S)$ and we have existence and uniqueness of IC's for X_H even for $k = 1$ [1].

3.1.3 Controlled Simple Mechanical Systems

We consider more general control systems that those considered by Lewis [3]. A *controlled second order system* is a triple $CS = (S, \mathcal{U}, F)$ such that (i) $S = TM$ where M is a smooth manifold; (ii)

$\mathcal{U}$ is a set of *open-loop* controls; and (iii) $F = \{F_u;\ u \in \mathcal{U}\}$ is a family of second-order Carathéodory vector fields on M. Let U be a Lebesgue-Borel measurable subset of $I \times \mathbb{R}^m$, where I is an interval, such that $U(t) = \{\bar{u} : (t, \bar{u}) \in U\}$ is nonempty for every $t \in I$. Denote the time domain of U by $\mathcal{TD}(U)$. Then each element of $\mathcal{U}$ is a function $u : I \to U$ where I is an interval and $\mathcal{TD}(U) = I$. Notice that U is not dependent on points $x \in S$. Contrary to the claim made by Lewis [3], the situation when U is dependent on $x \in S$ is more complicated and is not covered by the following theory [4]. We will make the following assumptions that will simplify matters: (a) F is a SOCVF on M such that $(q, v, t, u) \to F(q, v, t, u)$ is a map from $TM \times U$ to TTM with $F(q, v, t, u) \in T_{(q,v)}TM$ for all $(q, v, t, u) \in M \times U$. (b) every map $F(\cdot, \cdot, t, u)$ is continuous for each $(t, u) \in U$, (c) every map $F(q, v, \cdot, \cdot)$ is a Lebesgue-Borel measurable, where $q \in M$ and $v \in T_qM$.

A controlled trajectory of a system CS is a pair $\gamma = (c, u)$ where $u \in \mathcal{U}$ and (c, F_u) is a controlled curve. If Domain(c) is compact, then γ is called a *controlled arc*. $Ctraj(CS)$ (respectively, $Carc(CS)$) denotes the set of all controlled trajectories (respectively, controlled arcs) of CS. A trajectory (respectively, arc) of CS is a c such that $(c, u) \in Ctraj(CS)$ (respectively, $Carc(CS)$) for some $u \in \mathcal{U}$. Let CS be a controlled second order system and let $a, b \in \mathbb{R}$ with $a \leq b$. If $x_1 = (q_1, V_1)$, $x_2 = (q_2, V_2) \in TM$, we say that x_2 is *CS-reachable* from x_1 over $[a, b]$ if there exists a controlled arc γ of CS such that $c(a) = x_1$ and $c(b) = x_2$. x_2 is said to be CS-reachable from x_1 if it is CS-reachable from x_1 over $[a, b]$ for some a, b. Define $\mathcal{R}^{CS}(x)$ (respectively, $\mathcal{R}^{CS}_{[a,b]}(x)$) the CS-reachable set from x (respectively, the CS-reachable set from x over $[a, b]$).

A *Lagrangian* L for CS is a family $\{L_u;\ u \in \mathcal{U}\}$ of Carathéodory functions on S such that $\mathcal{TD}(L_u) = \mathcal{TD}(F_u)$ for all $u \in \mathcal{U}$. Given a Lagrangian L for CS, the cost functional $J^{L,CS} : Carc(CS) \to \mathbb{R}$ is defined by:

$$J^{L,CS}(\gamma) = \int_{\text{Domain}(\gamma)} L_u(c(t), t)\, dt \tag{18}$$

Following Sussmann [1], a controlled arc $\gamma = (c, u)$ is called *acceptable* for L if $|L_u(c(\cdot), \cdot)|$ is a locally integrable function. If γ is acceptable for L and has domain $[a, b]$, then the function $\rho_{\gamma,L}(t) = \int_a^t L_u(c(s), s)\, ds$ is the *running cost* along γ.

The augmented system CS^L associated with a second order system CS and a Lagrangian L is the system:

$$\dot{x} = F_u(x, t);\ \dot{x}^0 = L_u(x, t);\ \ u \in \mathcal{U}. \tag{19}$$

A trajectory $\hat{c}$ of the augmented system for a control $u \in \mathcal{U}$ is a pair (c, c^0) such that c is a trajectory of CS for u, (c, u) is acceptable for L, and c^0 is the running cost for (c, u). We refer to Sussmann

[1] for the definition of the substitution properties.

For each control system CS and Lagrangian L for CS, we can associate the *Hamiltonian*:

$$\mathcal{H}^{CS,L} \triangleq \{H^{F_u,L_u} : u \in \mathcal{U}\} \tag{20}$$

Due to our continuity assumptions on F, (leading to the so-called *classical time-varying system* defined by Sussmann) the optimal control $u^* \in \mathcal{U}$ will have the following strongly minimizing property along a curve (c, z) in T^*TM (by Prop. 6.1, Theorem 10.1 of [1]):

$$\mathcal{H}^{CS,L}(c(t), z(t), t, u^*) = \min_{\tilde{u} \in \mathcal{U}}\{\mathcal{H}^{CS,L}(c(t), z(t), t, \tilde{u})\} \quad \text{for a.e } t \in \mathcal{TD}(F_{\tilde{u}}) \cap \text{ Domain}(c) \tag{21}$$

An L-adjoint vector ζ along $\gamma = (c, u) \in Ctraj_{LIL}(CS)$ is an absolutely continuous section $\aleph(c, T^*TM)$ such that (c, ζ) is an integral curve of $X_{H^{F_u,L_u}}$. Hence the L-adjoint vectors are defined only along the trajectory c on TM. An L-adjoint vector field ζ along $\gamma = (c, u)$ is strongly minimizing for (CS, L), if u is strongly minimizing for (CS, L) along the curve (c, ζ) according to (21).

3.2 Pontryagin's Minimum Principle for a Simple Mechanical System

The PMP given in this section is a special case of Theorem 8.3 of [1] for Carathéodory second-order systems. The interesting aspect of our version is the choice of coordinates that facilitates the numerical solution of the optimal control.

Let M be a Riemannian manifold with coordinate charts $\{U_\alpha, \varphi_\alpha,\ \alpha \in I\}$. On a coordinate chart (U_α, Φ_α) of TM (see Section 2) let $\{e_1, \cdots, e_n\}$ be a frame of vector fields compatible with the metric. Similarly, on a coordinate chart (U_α, Ψ_α) of T^*M let $\{e^1, \cdots, e^n\}$ be a frame of co-vector fields so that $e^i(e_j) = \delta^i_j;\ 1 \le i, j \le n$. Let (q, v) denote coordinates on the chart $\Phi_\alpha(TU_\alpha)$. They form the state variables for a simple mechanical system on M. There is a natural connection defined on M called the Levi-Civita connection for which the metric is invariant [13]. Using the Levi-Civita connection on M, we can describe a controlled arc $\gamma = (c, u)$ (where $c(\cdot) = (q, v)(\cdot)$) for a controlled second order system (see 3.1.3 for the definition) $CS = (S, \mathcal{U}, F)$ as one on $HTM \oplus VTM$:

$$\dot{q}(t) = v(t) = v^i e_i, \quad \text{and} \quad \frac{Dv}{dt} = f(q(t), v(t), u(t), t) = f^i(q(t), v(t), u(t), t)e_i. \tag{22}$$

Consider the following assumptions that correspond to $A1, A3, A4$, and $OPT1 - OPT4$ of [1]:

MP1 $CS = (TM, \mathcal{U}, F)$ is a second-order control system as described in 3.1.3. The assumptions on F described there will apply to the functions f^i, $i = 1, \cdots, n$ in (22)

MP2 $a,\, b \in \mathbb{R}$, $a \leq b$, x_0, $x_f \in TM$, $c_* : [t_0, t_f] \to TM$ is a trajectory of CS with $c_*(t_0) = x_0$ and $c(t_f) = x_f$.

MP3 $u_* \in \mathcal{U}$ is such that $\gamma_* = (c_*, \zeta_*)$ is a LIL-controlled arc of CS.

MP4 Suppose N is a given submanifold of TM and $x_0 \notin N$ be a given point of TM. W is a neighborhood of (x_0, x_f) in $TM \times TM$ and $\varphi : W \to \mathbb{R}$ is a Lipschitz continuous function with $\mathbf{G} = \partial\varphi(x_f)$ – the Clarke generalized gradiant of φ at x_f.

MP5 L is a Lagrangian for CS and L_{u^*} is LIL.

The set $Carc^L(CS)$ denotes the set of controlled arcs $\gamma = (c, u)$ that are acceptable for L.

*R*emarks:

i. Notice that due to our assumption MP1, condition A2 of [1] for the augmented system (19) is automatically satisfied due to the fact that our time-varying second order system is a time-varying classical system according to [1]. Due to Theorem 10.1 of [1], weak minimization in the Maximum Principle for optimal control given in Theorem 8.3 in the same reference can be replaced by strong minimization.

ii. Unlike Sussmann [1] and Lewis [3] we will assume that x_0 is given to be the generalized position and velocity at time t_0. This reflects a trajectory planning problem in engineering applications.

iii. If $\varphi : W \to \mathbb{R}$ was a smooth function, then $\mathbf{G} = \{d\varphi(x_f)\}$.

We consider the following *f*ixed and final-time optimal control problems.

*P*roblem P1: Minimize the cost functional

$$J^{L,CS,\varphi}(b, \gamma) = \varphi(c(b)) + \int_a^b L_u(c(t), t)\, dt \tag{23}$$

in the set $\gamma = (c, u) \in Carc^L(CS)$, Domain$(c) = [a, b]$, and $c(b) \in N$.

*P*roblem P2: Minimize the cost functional

$$J^{L,CS,\varphi}(b, \gamma) = \varphi(c(b)) + \int_a^b L_u(c(t), t)\, dt \tag{24}$$

in the set (b, γ), such that $\gamma = (c, u) \in Carc^L(CS)$, Domain$(c) = [a, b]$, and $c(b) \in N$.

The theorem below gives the first order necessary conditions for the existence of the solution $\gamma_* = (c_*, u_*)$ to the optimal control problems. We need some notation before we can state the theorem.

- $(R^\flat(v,p_2)v,0)$ a one-form field along the curve (c,ζ) in H^*TM that satisfies $\langle R^\flat(v,p_2)v,\xi\rangle = \langle p_2, R(v,\xi)v\rangle$ for every $(\xi,0)\in H_{c(t)}TM$ for every $t\in\mathcal{TD}(c)$.

- We denote by $((d_qf)^*(p_2),0)$ a one-form field along the curve (c,ζ) in H^*TM that satisfies $\langle (d_qf)^*(p_2),\xi\rangle = \langle p_2, df_q(\xi)\rangle$, for every $(\xi,0)\in H_{c(t)}TM$ for every $t\in\mathcal{TD}(c)$. Similarly, denote by $(0,(d_vf)^*(p_2))$ a one-form field along the curve (c,ζ) in V^*TM that satisfies $\langle (d_vf)^*(p_2),\bar{w}\rangle = \langle p_2, d_vf(w)\rangle$, for every $(0,\bar{w})\in V_{c(t)}TM$ for every $t\in\mathcal{TD}(c)$.

- Let $C^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj}$ be the structure constants for the Jacobi-Lie bracket of vector fields on M. We rewrite the term $f^i\,\Gamma^j_{ki}\xi^k\,p_{2j}$ in terms of geometrically defined quantities. Denote $[\Gamma(f)]^*p_2 = p_{2j}\Gamma^j_{ik}f^ie^k$ and $[C(f)]^*p_2 = p_{2j}C^j_{ik}f^ie^k$, and then

$$f^i\,\Gamma^j_{ki}\xi^k\,p_{2j} = f^i\,(\Gamma^j_{ik} - C^j_{ik})\xi^k\,p_{2j} = \langle [\Gamma(f)]^*p_2 - [C(f)]^*p_2,\xi\rangle.$$

Theorem 3.1 *Suppose that assumptions MP1 - MP5 hold and suppose that γ_* is a solution of Problem P1. Then there exists a constant $\nu\geq 0$ and a $\nu\,L$ adjoint vector ζ along γ_* such that:*

i. *in local coordinates $\zeta(t) = ((p_1(t),0),(0,p_2(t)))\in H^*_{c(t)}TM\oplus V^*_{c(t)}TM$ (see Section 2 and Figures 3 -6 for the notation) satisfies:*

$$-\frac{Dp_1}{dt} = d_qf^*(p_2) + \nu\,d_qL - ([C(f)]^* - [\Gamma(f)]^*)p_2 + R^\flat(v,p_2)v \quad (25)$$
$$-\frac{Dp_2}{dt} = p_1 + d_vf^*(p_2) + d_vL. \quad (26)$$

ii. *ζ is strongly minimizing for $(CS,\nu\,L)$ along γ_* (see (21)).*

iii. *ζ satisfies the transversality condition: there exists a $\lambda\in\mathbf{G}$ such that $\langle\zeta(b)-\nu\lambda,\xi\rangle = 0$ for every $\xi\in T_{c_*(b)}N$.*

iv. *either $\zeta(b)\neq 0$ or $\nu>0$ (in other words, $\zeta(b)$ and ν cannot both be identically 0).*

v. *ζ and ν can be chosen so that the Hamiltonian $H^{F_{u^*},\nu\,L_{u^*}}$ (see (16)) is constant almost everywhere in $[a,b]$.*

For Problem P2, ζ and ν can be so chosen so that the constant value of the Hamiltonian is zero.

Proof: The theorem is a special case of Theorem 8.3 of Sussmann [1], and all of the conclusions except the first follow from this theorem. We will show that the adjoint vector $\zeta = (p_1,p_2)$ along γ_*

satisfies Equations (25 - 26), where (p_1, p_2) are local coordinates on $H^*_{c_*}TM \oplus V^*_{c_*}TM$. According to (16), the Hamiltonian function $H^{F_u,\nu L_u}$ is given in local coordinates by:

$$H^{F_u,L_u}(c,\zeta,t) = \langle \zeta, F(c,u,t)\rangle + \nu L(x,u,t) = \langle \zeta, \dot{c}\rangle + \nu L(x,u,t), \tag{27}$$

where $c(t) = (q(t), v(t))$ and $\dot{c}(t) = (v(t), F(x(t), v(t), u(t), t))$. Now we can choose $\zeta(t) = (\bar{p}_1, \bar{p}_2)(t)$ in local coordinates on $T^*_{c(t)}TM$ or $(p_1, p_2)(t)$ on $H^*_{c(t)}TM \oplus V^*_{c(t)}TM$. These choices are related by (10) as follows:

$$\begin{aligned} H^{F_u,L_u}(c,\zeta,t) &= \langle (\bar{p}_1, \bar{p}_2), (\dot{q}, \dot{v})\rangle + \nu L(x,u,t) \\ &= \langle p_1, \dot{q}\rangle + \langle p_2, \frac{Dv}{dt}\rangle + \nu L(x,u,t) \\ &= \langle p_1, \dot{q}\rangle + \langle p_2, f(c,u,t)\rangle + \nu L(x,u,t) \end{aligned} \tag{28}$$

At this point, obtaining the adjoint equations is simply a matter of applying the definition of a Hamiltonian vector field given in (17). Let $(\xi, w, \mu, \vartheta) \in T_{(q,v,p_1,p_2)}(H^*TM \oplus V^*TM)$. Then by (15) and (17) we have:

$$\Sigma_{(q,v,p_1,p_2)}((\dot{q}, \dot{v}, \dot{p}_1, \dot{p}_2), (\xi, w, \mu, \vartheta)) = dH^{F_u,L_u}_{(q,v,p_1,p_2)}(\xi, w, \mu, \vartheta) \tag{29}$$

where the LHS is:

$$\begin{aligned} \Sigma_{(q,v,p_1,p_2)}((\dot{q}, \dot{v}, \dot{p}_1, \dot{p}_2), (\xi, w, \mu, \vartheta)) &= \langle \dot{q}, (\mu_i - p_{1j}\Gamma^j_{ki}\xi^k)e^i\rangle - \langle \xi, (\dot{p}_{1i} - p_{1j}\Gamma^j_{ki}\dot{q}^k)e^i\rangle \\ &\quad - v^k p_{2i}\Omega^i_k(\dot{q}, \xi) + \langle (\dot{v}^i + \Gamma^i_{jk}v^k\dot{q}^j)e_i, (\vartheta_i - p_{2j}\Gamma^j_{ki}\xi^k)e^i\rangle \\ &\quad - \langle (w^i + \Gamma^i_{jk}\xi^j v^k)e_i, (\dot{p}_{2i} - p_{2j}\Gamma^j_{ki}\dot{q}^k)e^i\rangle \end{aligned}$$

The right hand side of (29)$dH^{F_u,L_u}_{(q,v,p_1,p_2)}(\xi, w, \mu, \vartheta)$ can be computed in two ways, both yielding the same result. Consider an equivalent class of LL curves in $V^*TM \oplus H^*TM$ with each curve $(c,\zeta) : (-\epsilon, \epsilon) \to V^*TM \oplus H^*TM$ satisfy: $(c,\zeta)(0) = (q, v, p_1, p_2)$ and $\frac{d}{ds}(c,\zeta)(0) = (\xi, w, \mu, \vartheta)$. Notice the slight abuse of notation, wherein we refer to $(c,\zeta)(0) = (q, v, p_1, p_2)(0)$ as (q, v, p_1, p_2). In the first method, one computes the derivative $\frac{d}{ds}H(0)$ using the techniques of Riemannian geometry. The other method is perhaps more interesting. The symplectic two-form Σ in (15) can be rewritten using the *exterior covariant differentials* [15] of $v = v^i e_i$; $p_1 = p_{1i}e^i$ and $p_2 = p_{2i}e^i$, as:

$$\Sigma = e^i \wedge (\nabla p_1)_i - p_{2i}(\nabla\nabla v)_i + (\nabla v)^i \wedge (\nabla p_2)_i, \tag{30}$$

where:

$$\begin{aligned} \nabla(v^i e_i) &= (dv^i + \Gamma^i{}_k v^k)\, e_i & \qquad \nabla(p_{1i}e^i) &= (dp_{1i} - p_{1j}\Gamma^j{}_i)\, e^i \\ \nabla(p_{2i}e^i) &= (dp_{2i} - p_{2j}\Gamma^j{}_i)\, e^i & \qquad \nabla\nabla(v^i e_i) &= v^k \Omega^i{}_k\, e_i \end{aligned}$$

The point of this discussion is that we must employ covariant derivatives while computing the right hand side of (29) as well! In the computation below, the evaluation is being done at $s = 0$.

$$
\begin{aligned}
dH^{F_u,L_u}_{(q,v,p_1,p_2)}(\xi, w, \mu, \vartheta) \;=\; & \nu \frac{\partial L}{\partial q^i} \left(\frac{\partial q}{\partial s}(0)\right)^i + \nu \frac{\partial L}{\partial v^i} \left(\nabla_{\frac{\partial q}{\partial s}} v\,(0)\right)^i + \langle \nabla_{\frac{\partial q}{\partial s}} p_1(0), v(0)\rangle \\
& + \langle p_1(0), \nabla_{\frac{\partial q}{\partial s}} v(0)\rangle + \langle \nabla_{\frac{\partial q}{\partial s}} p_2(0), f(q,v,u)(0)\rangle + \langle p_2(0), \frac{\partial f^i}{\partial q^j} \left(\frac{\partial q}{\partial s}(0)\right)^j e_i\rangle \\
& + \langle p_2(0), \frac{\partial f^i}{\partial v^j} \left(\nabla_{\frac{\partial q}{\partial s}} v(0)\right)^j e_i\rangle + p_{2j}(0) f^i(q,v,u)(0)\, \Gamma^j_{ki} \left(\frac{\partial q}{\partial s}(0)\right)^k
\end{aligned}
$$

After some simple manipulations, we get:

$$
\begin{aligned}
dH^{F_u,L_u}_{(q,v,p_1,p_2)}(\xi, w, \mu, \vartheta) \;=\; & \nu(d_q L)(\xi) + \nu(d_v L)(w + \Gamma_{jk}\xi^j v^k) + \langle \mu, v\rangle + \langle p_1, w\rangle \\
& + \langle \vartheta, f(q,v,u)\rangle + \langle p_2, (d_q f)(\xi)\rangle + \langle p_2, (d_v f)(w + \Gamma_{jk}\xi^j v^k)\rangle
\end{aligned}
$$

Now if we compare the terms on the left and right hand sides of (29) with μ, ϑ, w and ξ in that order, we get the desired Hamiltonian vector field:

$$
\begin{aligned}
\mu: \quad & \dot q = v \\
\vartheta: \quad & \tfrac{Dv}{dt} = f \\
w: \quad & -\tfrac{Dp_2}{dt} = p_1 + \nu d_v L + (d_v f)^* p_2 \\
\xi: \quad & -\tfrac{Dp_1}{dt} = \nu d_q L + (d_q f)^* p_2 + R^\flat(v, p_2)v + p_{2j}\Gamma^j{}_i f^i \\
& \qquad\;\; = \nu d_q L + (d_q f)^* p_2 + R^\flat(v, p_2)v - [C(f)]^* p_2 + [\Gamma(f)]^* p_2
\end{aligned}
\tag{31}
$$

□

In the special case of parallelizable Riemannian manifolds of which Lie Groups form a subset, the manifold TM is diffeomorphic to $M \times \mathbb{R}^n$ and hence the expressions for the first order necessary conditions are especially useful for numerical computations.

3.3 Cubic splines on Riemannian manifolds

Here we specialize the results of the previous section and recover the formula for cubic splines on Riemannian manifolds [16]. Let M be a parallelizable Riemannian manifold and let $q_0, q_1 \in M$, $v_0 \in T_{q_0}M$ and $v_1 \in T_{q_1}M$. Consider the problem: Minimize $J(u(\cdot)) = \frac{1}{2}\int_{t_0}^{t_f} \|u(t)\|^2 dt$ subject to: $\dot q(t) = v(t)$, $\frac{Dv}{dt} = u(t) = u^i(t) e_i(t)$, and boundary conditions $q(t_0) = q_0$; $q(t_f) = q_f$; $\dot q(t_0) = v_0$; $\dot q(t_f) = v_f$.

Thus we have $f(q, v, u) = u$ and the "flattening map" $^\flat : TM \to T^*M$ is the identity matrix, that is, the metric coefficients are given by $m_{rj} = \delta_{rj}$ for $r, j = 1, \cdots, n$. Therefore, there is an identification of vectors and co-vectors. The basic compatibility condition between the metric and connection coefficients is [14]: ${\Gamma^r}_i m_{rj} + {\Gamma^r}_j m_{ri} = 0 \quad i, j = 1, \cdots, n$. Using $m_{rj} = \delta_{rj}$, we get: ${\Gamma^j}_i + {\Gamma^i}_j = 0$ for $i,\, j = 1, \cdots, n$.

The Hamiltonian for this problem is

$$H^{F_u, L_u}(c, \zeta, t) = \langle p_1, v\rangle + \langle p_2, u(t)\rangle + \frac{\nu}{2}\, \|u(t)\|^2$$

If $\nu = 0$, then $p_2(t)$ must be zero for all t as the Hamiltonian is then a linear function of u. This implies that $p_1(t) = 0$ for all t by (26). This contradicts the PMP's assertion that both ν and ζ cannot both be 0. Hence $\nu > 0$ and it can be chosen to be 1 (see [1]).

The PMP asserts the existence of one-form sections $(p_1, p_2)(t)$ such that: $\frac{Dp_1}{dt} = -R^\flat(v, p_2)v + (\Gamma^*(u) - C^*(u))p_2$; $\frac{Dp_2}{dt} = -p_1$ and $u = -p_2$, where we have used the identification of vectors and co-vectors in the last equation. Thus $\frac{D^2v}{dt^2} = \frac{Du}{dt} = -\frac{Dp_2}{dt} = p_1$ which implies $\frac{D^3v}{dt^3} = \frac{Dp_1}{dt} = -R(v, p_2)v - (\Gamma^*(u) - C^*(u))\nu$, where we have again the identification of vectors and co-vectors. Now $((\Gamma^*(u) - C^*(u))p_2)v = p_{2i}({\Gamma^i}_j(u) - {C^i}_j(u))\, v^j = p_{2i}\, \Gamma^i_{jk}\, p_2^k\, v^j$ where $p_2^k = p_{2k}\ \forall k = 1, \cdots, n$. Therefore, $((\Gamma^*(u) - C^*(u))p_2)v = {\Gamma^i}_k(v)\, p_2^k\, p_{2i} = 0$ for all $v \in \Psi(M)$, because ${\Gamma^i}_k = -{\Gamma^k}_i$ by the basic compatibility condition for the connection with the metric. Therefore, $\frac{D^3v}{dt} + R(\frac{Dv}{dt}, v)v = 0$, which is the equation for a cubic spline that was obtained by Noakes, Heinzinger and Paden [16].

3.4 The Adjoint Equation

In this section, we attempt specialize our results to those in Lewis [3], by considering particular functions f in (22); L in (23) and (24); while using frames that are compatible with a Riemannian metric on M. Lewis [3] expresses his results in terms of the torsion, while we do not have torsion terms in (25) and (26). Our frames are parallel translated and hence are torsion free while the structure constants are not zero (see (25)). In [3], the structure constants are zero due to the use of coordinate frames while the torsion is not zero. Simply setting the torsion terms to zero in the adjoint equation of [3] results in a equation that differs from ours by one term. We claim here that this term is actually present in the adjoint equation and can be rewritten so that our results and [3] are identical.

For each $u \in \mathcal{U}$, let A_u be a symmetric $(0, r)$ tensor field on M [3]. Hence, $A_u : \aleph^r(M, TM) \to \mathbb{R}$. Let $l : \mathbb{R} \to \mathbb{R}$ be a LIL function. Consider the Lagrangian function $L : \aleph^r(M, TM) \times \mathcal{U} \to \mathbb{R}$ defined by $L = l \circ A_u$. Furthermore, for each $u \in \mathcal{U}$, let $f_u : TM \to TM$ be defined by: $f_u(q, v) = u^i(t)\, Y_i(q(t))$, where $Y_i \in \aleph(M, TM)$ and i lies in some finite index set. Thus, f_u does not explicitly depend on the velocity variable v.

We only need to consider the equations for the co-states (25) and (26). Following [3], for each $u \in \mathcal{U}$ define: $\widehat{A}_u : \aleph^{r-1}(M, TM) \to \aleph(M, T^*M)$, by: $\langle \widehat{A}_u(v(t), \cdots, v(t)), X \rangle = A_u(X, v(t), \cdots, v(t))$, where $X \in \aleph(M, TM)$. Then we have: $d_v L = \frac{\partial L}{\partial v^i} e^i = r\, l'(A_u(v(t), \cdots, v(t))\, \widehat{A}_u(v(t), \cdots, v(t))$, so that:

$$-\frac{Dp_2}{dt} = p_1 + \nu d_v L = p_1 + \nu\, r\, l'(A_u(v(t), \cdots, v(t)))\, \widehat{A}_u(v(t), \cdots, v(t)). \tag{32}$$

Next, we have: $d_q L = r\, l'(A_u(v(t), \cdots, v(t)))\, \nabla A_u(v(t), \cdots, v(t))$, and $(d_q f)^* p_2 = u^i\, (\nabla Y_i)^*(p_2)$. Here, $(\nabla Y_i)^*(p_2)$ is defined by $\langle (\nabla Y_i)^*(p_2), X \rangle = \langle p_2, \nabla_X\, Y_i \rangle$ where $X \in \aleph(M, TM)$, and $\nabla A_u(v(t), \cdots, v(t))$ is defined similarly.

$$\begin{aligned} -\frac{Dp_1}{dt} &= \nu\, r\, l'(A_u(v(t), \cdots, v(t))\, \nabla A_u(v(t), \cdots, v(t)) + u^i\, (\nabla Y_i)^*(p_2) + R^\flat(v, p_2)v \\ &\quad + p_{2j} \Gamma^j_{\ i}\, Y^i_k(q(t)) u^k(t) \end{aligned} \tag{33}$$

We claim that the last term in the above equation is the term $u^k(t)\, \frac{1}{2} T^*(p_2(t), Y_k(t))$ found in [3]. The reason for this is that Lewis considers the torsion to be given by: $T^j_{ki} = \Gamma^j_{ki} - \Gamma^j_{ik}$ which means that the structure constants are 0, which is the case when one uses coordinate frames. While using coordinate frames, one has: $T^j_{ki} = -T^j_{ik}$ and hence one can take Γ^j_{ki} to be: $\Gamma^j_{ki} = \frac{1}{2} T^j_{ki}$ by anti-symmetry in the k and i indices. Now following Lewis, define $T^* : \aleph(M, T^*M) \times \aleph(M, TM) \to \aleph(M, T^*M)$ to be the operator defined by: $\langle T^*(p_2, X), Z \rangle = \langle p_2, T(Z, X) \rangle$, where $X,\ Z \in \aleph(M, TM)$. Then we have for a.e. $t \in [a, b]$:

$$\begin{aligned} \langle u^k(t)\, \frac{1}{2} T^*(p_2(t), Y_k(t)), Z \rangle &= \langle p_2(t), \frac{1}{2}\, T(Z, Y_k(t)) u^k(t) \rangle \\ &= \langle p_2(t), \Gamma^j_{\ i}(Z) Y_k(t) u^k(t) \rangle \\ &= \langle p_{2j} \Gamma^j_{\ i}\, Y^i_k(q(t)) u^k(t), Z \rangle, \end{aligned}$$

for every $Z \in \aleph(M, TM)$ as desired. Note here that the term in question was also obtained in the form given in (25) using entirely different methods [2]. Equations (32) and (25) when combined together form the equation that Lewis terms *the adjoint equation.*

4 Conclusion

Sussmann [1] extended Pontryagin's Minimum Principle for optimal control problems on manifolds. This result did not use any additional structure on manifolds other than Lie brackets. As a result the theorem was obtained abstractly in terms of the Hamiltonian vector field of a Hamiltonian function. When applying this result in applications, one has to make a suitable choice of co-ordinates. In problems arising in robotics and engineering, the manifold Q is the tangent bundle TM of a Riemannian manifold M. Sussmann's result was specialized to control-affine systems on Riemannian manifolds with an affine connection by Lewis [3]. Lewis obtained an explicit expression for the Hamiltonian vector field by using the fact that it can thought of as the cotangent lift of the geodesic spray. In this paper, we have specialized the results of Sussmann [1] by using a different technique. We used the fact that the Hamiltonian vector field on the cotangent bundle of a manifold Q can be derived using the natural symplectic two-form that exists on it. Using this procedure, obtained explicit expressions for the vector fields of the adjoint system. These expressions turn out to be the same as those obtained in our earlier work [2] where we employed calculus of variations. The utility of this approach for numerical computations was demonstrated in [2].

References

[1] H. Sussmann, "An introduction to the coordinate-free maximum principle," in *Geometry of Feedback and Optimal Control* (B. Jakubczyk and W. Respondek, eds.), pp. 463–557, Marcel Dekker, New York, 1997.

[2] R. V. Iyer, R. Holsapple, and D. Doman, "Optimal Control Problems on Parallelizable Riemannian Manifolds: Theory and Application." in press, ESAIM: Control, Optimisation and Calculus of Variations, 2004.

[3] A. Lewis, "The geometry of the maximum principle for affine connection control systems." available online at http://penelope.mast.queensu.ca/~andrew/cgi-bin/pslist.cgi?papers.db, 2001.

[4] R. Vinter, *Optimal Control.* Birkhauser, 2000.

[5] V. Jurdejevic, *Geometric Control Theory.* Cambridge Studies in Advanced Mathematics, 1997.

[6] P. S. Krishnaprasad, "Optimal control and Poisson reduction," TR 93-87, Institute for Systems Research, University of Maryland, 1993.

[7] W. S. Koon and J. E. Marsden, "Optimal control for holonomic and nonholonomic mechanical systems with symmetry and lagrangian reduction," *SIAM Journal on Control and Optimization*, vol. 35, no. 3, pp. 901–929, 1997.

[8] R. W. R. Darling, *Differential Forms and Connections.* Cambridge University Press, 1994.

[9] I. Kolář, P. W. Michor, and J. Slovák, *Natural Operations in Differential Geometry.* Springer Verlag, 1993.

[10] P. Libermann and C.-M. Marle, *Symplectic Geometry and Analytical Mechanics.* IMA, 1987.

[11] R. Abraham and J. E. Marsden, *Foundations of Mechanics.* Perseus Books, 1985.

[12] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry.* Springer Verlag, 1994.

[13] M. P. D. Carmo, *Riemannian Geometry.* Birkhauser, 1993.

[14] W. M. Boothby, *An introduction to Differential Geometry and Riemannian Manifolds.* Academic Press, 1975.

[15] T. Frankel, *Geometry of Physics.* Cambridge Universit Press, 1997.

[16] L. Noakes, G. Heinzinger, and B. Paden, "Cubic splines on curved spaces," *IMA Journal of Mathematical Control and Information*, vol. 6, pp. 465–473, 1989.