# The geometry of the maximum principle for affine connection control systems*

Andrew D. Lewis†

2000/05/17
Last updated: 2001/01/08

**Abstract**

The maximum principle of Pontryagin is applied to systems where the drift vector field is the geodesic spray corresponding to an affine connection. The result is a second-order differential equation whose right-hand side is the "adjoint Jacobi equation." This latter relates the Hamiltonian geometry of the maximum principle to the affine differential geometry of the system's affine connection. The resulting version of the maximum principle is then be applied to the situation where the cost function is the norm squared of the input force.

**Keywords.** optimal control, mechanical systems, affine connections.

**AMS Subject Classifications (2020).** 49K15, 49Q20, 53B05, 70Q05, 93B27.

## Contents



*Appears as part of internet supplement to *Geometric Control of Mechanical Systems*, Francesco Bullo and Andrew D. Lewis, Springer-Verlag, New York-Heidelberg-Berlin, 2004, Number 49 in Texts in Applied Mathematics, ISBN 0-387-22195-6.

†Professor, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
Email: `andrew.lewis@queensu.ca`, URL: `http://www.mast.queensu.ca/~andrew/`

Research partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

## 1. Introduction

When one undertakes an investigation of a class of physical systems, an important first step is to obtain a useful general setting from which to launch the investigation. This paper forms part of an investigation of a class of physical systems which are called "simple mechanical control systems." These systems are control systems whose uncontrolled dynamics are Lagrangian with a "kinetic energy minus potential energy" Lagrangian function. The

kinetic energy is defined by a Riemannian metric which provides the Levi-Civita connection as a possible tool for studying the control problem. An investigation of "configuration controllability" for simple mechanical control systems is undertaken by Lewis and Murray [1997]. There the rôle of the affine connection is revealed as it applies to the control problem. The applicability of the analysis of Lewis and Murray is broadened by the realisation that the dynamics of systems with nonholonomic constraints, linear in the velocities, are in fact geodesics of a certain class of affine connections. This seems to first appear in the literature in the paper of Synge [1928]. The control theoretic setting is described in [Lewis 2000].[1] This suggests that an appropriate general setting for simple mechanical control systems, including those with constraints, is one where the uncontrolled dynamics are the geodesic equations of an arbitrary affine connection. In Section 7 we dub such systems "affine connection control systems."

In this paper we look at the optimal control problem for control systems in this setting. The notion of optimality is a natural one to study in mechanics—the Euler-Lagrange equations themselves are derived from a variational principal. Indeed, mechanics is justifiably considered by some to be a special case of optimal control. But a basic investigation of mechanical optimal control systems has yet to be undertaken. Let us review the existing literature. Koon and Marsden [1997] and Ostrowski, Desai, and Kumar [1999] look at optimal control for a class of nonholonomic systems with symmetry, especially the snakeboard in the latter paper. Also in the latter paper, the objective is to optimally select "gaits" in order to perform a prescribed manoeuvre. For the systems considered by these authors, it is natural to posit a cost function which is quadratic in the inputs, and so the controlled extremals will be smooth. The problem of force minimisation is also investigated by Crouch and Silva Leite [1991] and Noakes, Heinzinger, and Paden [1989]. In both of these papers, full actuation is assumed. The former work appears in more detail in [Crouch and Silva Leite 1995], with the further addition of state constraints. Other recent work includes the papers [Silva Leite, Camarinha, and Crouch 2000] and [Crouch, Silva Leite, and Camarinha 2000]. In the former paper, underactuation and state constraints are allowed. In the latter paper, a local description of the Hamiltonian structure of the variational problems of [Crouch and Silva Leite 1991] and [Noakes, Heinzinger, and Paden 1989] is explored. All of the work we cite above has the feature that there are no control constraints, so that a variational approach is possible, i.e., the full generality of the maximum principle is not needed. Here we take an approach which includes all of the above work by stating a general version of the maximum principle for affine connection control systems. This version of the maximum principle extracts the geometry inherent in such problems. We should mention, however, that the consequences of symmetry as explored in certain of the above papers does not appear as part of our analysis here, but may be an interesting topic for future research. Many of the formulae we present here simplify for cases when the configuration space is a Lie group, for example.

The main contribution of the paper may be roughly summarised as follows. The restatement we give for the maximum principle for affine connection control systems has the feature of being both intrinsic and global. The former ensures that the results are independent of coordinate system, and this is certainly a desirable property. Note that this

[1] The author wishes to acknowledge the paper of Bloch and Crouch [1995] for providing the impetus for the work [Lewis 2000].

property of being intrinsic is *not* a feature of the statement of the maximum principle in Euclidean space. More precisely, that part of the Hamiltonian equations of the maximum principle which is typically called the "adjoint equation" is not coordinate independent. In a general setting, this is addressed by Sussmann [1998]. Sussmann's approach provides a coordinate-invariant "adjoint equation," but his approach has the drawback that the adjoint equation is provided on an extremal-by-extremal basis. That is to say, given a candidate extremal, one may provide an adjoint equation associated with it. His approach does not provide a global adjoint equation which may be applied along any extremal—indeed, such an equation simply does not exist without additional structure. However, it is important to note that Sussmann's approach is, as far as the author knows, the first systematic acknowledgement that there is a need for an intrinsic adjoint equation. Owing to the extra structure in our problem, notably the structure of an affine connection, we are able to provide a *global* and intrinsic adjoint equation. In essence, we are able to globally "decouple" the Hamiltonian equations of the maximum principle into the control equation and a separate equation describing the evolution of the adjoint vector. It is this that we mean when we say our formulation of the maximum principle is global and intrinsic.

Let us give a rough outline of the paper. The optimal control problem formulated is one which properly lives on the state space, the tangent bundle of the configuration space, and not the configuration space itself. Nonetheless, since the problem data is all described on the configuration space, it is reasonable to anticipate that conditions for optimality may be formulated on configuration space. This is accomplished with an investigation of the tangent bundle geometry which can be attached to an affine connection. In particular, we review how one may define Ehresmann connections on various bundles. In the splittings defined by these connections, we determine that the Hamiltonian vector field provided by the maximum principle produces an equation which evolves on configuration space. This equation we dub the "adjoint Jacobi equation," and it is related to the Jacobi equation in affine differential geometry. To facilitate a full appreciation of the relationship between the Jacobi equation, which describes geodesic variation, and the adjoint Jacobi equation, the two equations are developed in parallel and in similar settings. In consequence, we devote a the first part of the paper to describing the geometry which produces these equations. In doing so, we carefully organise and provide alternate descriptions of certain previously known aspects of the tangent bundle geometry for manifolds with an affine connection (see Yano and Ishihara [1973]). The results here are made somewhat less transparent by our consideration of general affine connections, especially those with torsion. We confine this geometric discussion to Part I

After reviewing the relevant aspects of tangent bundle geometry, in Part II we look carefully at the optimal control problems we wish to study. Once we have a complete understanding of the adjoint Jacobi equation, it is a comparatively simple matter to use it to provide a version of the maximum principle for manifolds with an affine connection. First we formulate precisely a version of the maximum principle which is valid for general control affine systems (note that "control affine systems" and "affine connection control systems" differ!) on manifolds as developed by Sussmann [1998]. This development includes a nonstandard means of describing control constraints. Fortunately, the generality of Sussmann's results allows us to use his results in our setting. An appropriate adaptation of these results, using the development of Section 3, produces the desired necessary conditions for optimality, formulated on configuration space (Theorem 7.7). We then investigate

the optimal control problem where one minimises the force expended by the inputs along the trajectory. In this case the Hamiltonian to be maximised is convex in control, and so we may explicitly obtain the optimal controls from the maximum principle, at least for nonsingular extremals.

One may, of course, proceed to investigate other optimal control problems than the one we look at in Section 8. An investigation of time-optimal control for this class of systems has been initiated by Chyba, Leonard, and Sontag [2000], although the geometry of their results remains to be flushed out. It also remains to undertake a systematic investigation of matters like, e.g., classification of regular and singular extremals, and determining whether there are abnormal minimisers. Owing to the length of the paper, it is not possible to undertake a detailed investigation of specific physical examples. We refer to [Coombs 2000] for a discussion of the time-optimal problem for a few simple example. An interesting avenue for further work arises from the recent work of Bullo and Lynch [2001]. Using what they call "decoupling vector fields," they propose a concatenation of optimal arcs, with the concatenation not necessarily being itself optimal. It is perhaps interesting to investigate how this scheme interacts with the geometry in our paper.

## 2. Definitions and notation

Along with the basic definitions provided here, we give a list of symbols at the end of the paper.

When we write $A \subset S$ we mean to allow the possibility that $A = S$. If $S$ is a set, $\mathrm{id}_S \colon S \to S$ is the identity map on $S$. If $S$ is a set, $\mathbf{2}^S$ denotes the set of subsets of $S$.

The letter $I$ will denote a nonempty interval in $\mathbb{R}$, and unless otherwise stated we allow $I$ to be semi-infinite or infinite. Also, unless otherwise stated, it may be either open or closed at either of its endpoints.

If $V$ is a $\mathbb{R}$-vector space, $V^*$ denotes its dual, and we write $\langle \alpha; v \rangle$ to denote the natural pairing of $\alpha \in V^*$ with $v \in V$. If $S \subset V$ then we denote

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

As concerns basic differential geometric notation, we shall essentially follow Abraham, Marsden, and Ratiu [1988]. All manifolds will be finite-dimensional, separable, and $C^\infty$. Similarly, unless otherwise stated, all objects—functions, vector fields, etc. —will be assumed to be of class $C^\infty$. The tangent bundle of a manifold $M$ is denoted $\pi_{TM} \colon TM \to M$ and the cotangent bundle by $\pi_{T^*M} \colon T^*M \to M$. $C^\infty(M)$ denotes the set of smooth functions on $M$, and $\Gamma^\infty(E)$ denotes the smooth sections of a vector bundle $\pi \colon E \to B$. The set of tensors of type $(r, s)$ of a vector bundle $\pi \colon E \to B$ is denoted $\mathrm{T}^r_s(E)$. If $\phi \colon M \to N$ is a smooth map, $T\phi \colon TM \to TN$ denotes its derivative, and $T_x\phi$ is the restriction of $T\phi$ to the tangent space $T_xM$. The Lie derivative with respect to a vector field $X$ is denoted $\mathscr{L}_X$.

We shall work with local coordinates on manifolds $TQ$, $T^*Q$, $TTQ$, and $T^*TQ$ (with $Q$ itself a manifold), so it will be desirable to have a consistent way of writing such coordinates. If $(q^1, \dots, q^n)$ are coordinates for $Q$, natural coordinates for $TQ$ are denoted $(q^1, \dots, q^n, v^1, \dots, v^n)$. We shall often denote these as $(q, v)$ for brevity. Using this abbreviated notation, natural coordinates for $T^*Q$ are denoted $(q, p)$, and coordinates for $TTQ$

and $T^*TQ$ are denoted $((q,v),(u,w))$ and $((q,v),(\alpha,\beta))$, respectively. In each of the latter two cases, the first pair of coordinates refers to coordinates for the base space which is $TQ$, and the second pair of coordinates are the fibre coordinates.

Following Aliprantis and Border [2006] and Sussmann [1998], a ***Carathéodory function*** on a manifold $M$ is a map $\phi\colon I\times M\to\mathbb{R}$ where

CF1. $I\subset\mathbb{R}$ is an interval;

CF2. the function $x\mapsto\phi(t,x)$ is continuous for each $t\in I$;

CF3. the function $t\mapsto\phi(t,x)$ is measurable for each $x\in M$.

A Carathéodory function $\phi\colon I\times M\to\mathbb{R}$ is ***locally integrally bounded*** (LIB) if for every compact subset $K\subset M$ there exists a nonnegative, locally Lebesgue integrable function $\psi\colon I\to\mathbb{R}$ with the property that $|\phi(t,x)|\le\psi(t)$ for every $(t,x)\in I\times K$. A Carathéodory function $\phi\colon I\times M\to\mathbb{R}$ is ***of class $C^k$*** if the function $x\mapsto\phi(t,x)$ is of class $C^k$ for each $t\in I$. A Carathéodory function $\phi\colon I\times M\to\mathbb{R}$ is ***locally integrable of class $C^k$*** (LIC$^k$) if it is of class $C^k$ and if the function $(t,x)\mapsto X_1\cdots X_k\phi(t,x)$ is LIB for every collection $\{X_1,\dots,X_k\}$ of $C^\infty$ vector fields on $M$. We say $\phi$ is ***locally integrable of class $C^\infty$*** (LIC$^\infty$) if it is LIC$^k$ for every integer $k\ge 0$.

A ***Carathéodory vector field*** is a map $X\colon I\times M\to TM$ where

CVF1. $I\subset\mathbb{R}$ is an interval;

CVF2. $X(t,x)\in T_xM$ for each $(t,x)\in I\times M$;

CVF3. $(t,x)\mapsto\mathscr{L}_X\phi(t,x)$ is a Carathéodory function for each $\phi\in C^\infty(M)$.

A Carathéodory vector field $X\colon I\times M\to TM$ is LIC$^k$ if $(t,x)\mapsto X(t,x)\phi$ is an LIC$^k$ function on $M$ for $\phi\in C^\infty(M)$. A ***second-order Carathéodory vector field*** on $M$ is a Carathéodory vector field $X\colon I\times TM\to TTM$ on $TM$ with the property that $T\pi_{TM}\circ X(t,v_x)=v_x$ for every $(t,v_x)\in I\times TM$.

Recall that a function $f\colon I\to\mathbb{R}$, where $I\subset\mathbb{R}$ is an interval, is ***locally absolutely continuous*** (LAC) if it is almost everywhere differentiable and if for every compact subinterval $[a,b]\subset I$ there exists a bounded measurable function $g\colon[a,b]\to\mathbb{R}$ with the property that

$$(f|[a,b])(t)=\int_a^t g(s)\mathrm{d}s.$$

A function $f\colon I\to\mathbb{R}$ is ***locally absolutely differentiable*** (LAD) if for every compact subinterval $[a,b]\subset I$ there exists an absolutely continuous function $g\colon[a,b]\to\mathbb{R}$ with the property that

$$(f'|[a,b])(t)=\int_a^t g(s)\mathrm{d}s.$$

It is evident that a locally absolutely differentiable function is of class $C^1$ and that it is almost everywhere twice differentiable.

A ***curve*** on a manifold $M$ is a continuous map $c\colon I\to M$ where $I$ is an interval. A curve $c$ is LAC or LAD if $\phi\circ c$ is respectively LAC or LAD for every $\phi\in C^\infty(M)$. Given a Carathéodory vector field $X\colon I\times M\to TM$, an ***integral curve*** for $X$ is an LAC curve

$c\colon J \to M$ with the property that $c'(t) = X(t, c(t))$ for almost all $t \in J$ where $J$ is a subinterval of $I$. Recall that $\mathrm{LIC}^k$, $k \geq 1$, vector fields possess a maximal integral curve with initial condition $x$ at time $t = a$. If $X\colon I \times TM \to TTM$ is a second-order $\mathrm{LIC}^\infty$ Carathéodory vector field then its integral curves are LAC curves $\sigma\colon J \to TM$. Therefore there exists a LAD curve $c\colon J \to M$ which has the property that $c'(t) = \sigma(t)$ for every $t \in J$. Let $\pi\colon E \to B$ be a vector bundle. If $c\colon I \to B$ is a LAC or LAD curve, a LAC or LAD ***section of $\pi$ along $c$*** is a LAC or LAD map $\chi\colon I \to E$ with the property that $\pi \circ \chi(t) = c(t)$ for $t \in I$. Note that a section of $\pi$ along $c$ cannot have stronger smoothness properties than does $c$. We shall primarily be interested in vector fields and one-form fields along curves.

# Part I. Tangent Bundles and Affine Connections

## 3. Tangent bundle geometry

We will be dealing with the maximum principle as it applies to systems whose drift vector field is a second-order vector field, which is thus a vector field on the tangent bundle $TQ$ of a configuration manifold $Q$. In this section we introduce some of the necessary tangent bundle geometry which we shall use in our treatment of the optimal control problem. Furthermore, some of these notions are useful for understanding a geometric form of the maximum principle, even as it applies to general systems which are affine in the controls. Most of the constructions we make here are described by Yano and Ishihara [1973], at least in a time-independent setting.

The reader will observe in this section an alternation between the use of the letters $M$ and $Q$ to denote a generic manifold. There is some method behind this. We shall have occasion to use structure which, in the geometric constructions of Section 5, might appear on either $Q$ or its tangent bundle $TQ$. Such structures we will denote here as occurring on $M$. That is, when you see an $M$ in this section, it might refer to either $Q$ or $TQ$ in later sections. We shall suppose both $M$ and $Q$ to be $n$-dimensional in this section.

**3.1. The tangent lift of a vector field.** Let $X\colon I \times M \to TM$ be a $\mathrm{LIC}^\infty$ vector field on a manifold $M$. For $t \in I$ let $s \mapsto F_{t,s}$ be the one-parameter family of diffeomorphisms defining the flow of the $C^\infty$ vector field $x \mapsto X(t, x)$. Thus

$$\frac{\mathrm{d}}{\mathrm{d}s} F_{t,s}(x) = X(t, F_{t,s}(x)) \quad \text{and} \quad F_{t,0}(x) = x.$$

In particular, $F_{t,s}$ is *not* the flow of the time-dependent vector field $X$. We define a $\mathrm{LIC}^\infty$ vector field $X^T$ on $TM$ by

$$X^T(t, v_x) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} T_x F_{t,s}(v_x).$$

One may verify in coordinates that

$$X^T = X^i \frac{\partial}{\partial x^i} + \frac{\partial X^i}{\partial x^j} v^j \frac{\partial}{\partial v^i}. \tag{3.1}$$

From this coordinate expression, we may immediately assert a few useful facts.

**3.1 Remarks:** 1. Note that $X^T$ is a linear vector field on $TM$ (see Section 5.3). That is, $X^T$ is projectable and $X^T\colon I\times TM\to TTM$ is a vector bundle mapping.

2. Since $X^T$ is projectable and projects to $X$, if $t\mapsto v(t)$ is an integral curve for $X^T$, then this curve covers the curve $t\mapsto \pi_{TM}\circ v(t)$, and this latter curve is further an integral curve for $X$. Thus integral curves for $X^T$ may be thought of as vector fields along integral curves for $X$.

3. Let $x\in M$ and let $c$ be the integral curve for $X$ with initial condition $x$ at time $t=a$. Let $v_{1,x},v_{2,x}\in T_xM$ with $c_1^T$ and $c_2^T$ the integral curves for $X^T$ with initial conditions $v_{1,x}$ and $v_{2,x}$, respectively, at time $t=a$. Then $t\mapsto \alpha_1\,c_1^T(t)+\alpha_2\,c_2^T(t)$ is the integral curve for $X^T$ with initial condition $\alpha_1\,v_{1,x}+\alpha_2\,v_{2,x}$, for $\alpha_1,\alpha_2\in\mathbb{R}$. That is to say, the family of integral curves for $X^T$ which cover $c$ is a $\dim(M)$-dimensional vector space.

4. One may think of $X^T$ as the "linearisation" of $X$ in the following sense. Let $c\colon I\to M$ be the integral curve of $X$ through $x\in M$ at time $t=a$ and let $c^T\colon I\to TM$ be the integral curve of $X^T$ with initial condition $v_x\in T_xM$ at time $t=a$. Choose a smooth one-parameter family of deformations $\sigma\colon I\times[-\epsilon,\epsilon]\to$ of $c$ with the following properties:

   (a) $s\mapsto\sigma(t,s)$ is differentiable for $t\in I$;

   (b) for $s\in[-\epsilon,\epsilon]$, $t\mapsto\sigma(t,s)$ is the integral curve of $X$ through $\sigma(a,s)$ at time $t=a$;

   (c) $\sigma(t,0)=c(t)$ for $t\in I$;

   (d) $v_x=\frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}\sigma(0,s)$.

   We then have $c^T(t)=\frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0}\sigma(t,s)$. Thus $X^T(v_x)$ measures the "variation" of solutions of $X$ when perturbed by initial conditions lying in the direction of $v_x$. In cases where $M$ has additional structure, as we shall see, we can make more precise statements about the meaning of $X^T$.

The lift $X^T$ of $X$ from $M$ to $TM$ we call the ***tangent lift*** of $X$. There is another canonical way to lift a vector field to the tangent bundle. For a $\mathrm{LIC}^\infty$ vector field $X$ on $M$ define the ***vertical lift*** of $X$ to be the $\mathrm{LIC}^\infty$ vector field, $\mathrm{vlft}(X)$, on $TM$ given by

$$\mathrm{vlft}(X)(t,v_x)=\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}(v_x+sX(t,x)).$$

If $X=X^i\frac{\partial}{\partial x^i}$ in a local chart, then $\mathrm{vlft}(X)=X^i\frac{\partial}{\partial v^i}$ in the associated natural coordinates for $TM$.

**3.2. The cotangent lift of a vector field.** There is also a cotangent version of $X^T$ which we may define in a natural way. If $X$ is a $\mathrm{LIC}^\infty$ vector field on $M$, we define a $\mathrm{LIC}^\infty$ vector field $X^{T^*}$ on $T^*M$ by

$$X^{T^*}(t,\alpha_x)=\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}T_x^*F_{t,-s}(\alpha_x).$$

In coordinates we have

$$X^{T^*}=X^i\frac{\partial}{\partial x^i}-\frac{\partial X^j}{\partial x^i}p_j\frac{\partial}{\partial p_i}. \tag{3.2}$$

As was the case with $X^T$, we may make some immediate useful remarks about the meaning of $X^{T^*}$.

**3.2 Remarks:** 1. $X^{T^*}$ is the LIC$^\infty$ Hamiltonian vector field (with respect to the natural symplectic structure on $T^*M$) corresponding to the LIC$^\infty$ Hamiltonian $H_X\colon (t,\alpha_x) \mapsto \langle \alpha_x; X(t,x)\rangle$. (We adopt the convention that the natural symplectic form on $T^*M$ is given by $\omega_0 = -\boldsymbol{d}\theta_0$ where $\theta_0$ is the Liouville one-form. In coordinates, $\omega_0 = \mathrm{d}q^i \wedge \mathrm{d}p_i$.)

2. Note that $X^{T^*}$ is a linear vector field. That is, $X^{T^*}$ is projectable and $X^{T^*}\colon I \times T^*M \to TT^*M$ is a vector bundle mapping.
3. If $t \mapsto \alpha(t)$ is an integral curve for $X^{T^*}$, then this curve covers the curve $t \mapsto \pi_{T^*M} \circ \alpha(t)$, and this latter curve is further an integral curve for $X$. Thus one may regard integral curves of $X^{T^*}$ as one-form fields along integral curves of $X$.
4. If $c\colon I \to M$ is the integral curve for $X$ with initial condition $x \in M$ at time $t = a \in I$, then the integral curves of $X^{T^*}$ with initial conditions in $T^*_xM$ form a $\dim(M)$-dimensional vector space which is naturally isomorphic to $T^*_xM$ (in a manner entirely analogous to that described for $X^T$ in Remark 3.1–3).

**3.3. Joint properties of the tangent and cotangent lift.** By the very virtue of their definitions, together $X^T$ and $X^{T^*}$ should possess some properties. To formulate one of these common properties requires some small effort. Let $TM \oplus T^*M$ be the Whitney sum of $TM$ and $T^*M$. As a manifold, this may be regarded as an embedded submanifold of $TM \times T^*M$ by $v_x \oplus \alpha_x \mapsto (v_x, \alpha_x)$. In this way we identify $TM \oplus T^*M$ with the fibre product $TM \times_M T^*M$. If $X$ is a LIC$^\infty$ vector field on $M$, we define a LIC$^\infty$ vector field $X^T \times X^{T^*}$ on $TM \times T^*M$ by

$$X^T \times X^{T^*}(t,v,\alpha) = (X^T(t,v), X^{T^*}(t,\alpha)).$$

Note that in this definition we do not require that $\pi_{TM}(v) = \pi_{T^*M}(\alpha)$.

**3.3 Lemma:** *$X^T \times X^{T^*}$ is tangent to $TM \oplus T^*M$.*

Proof: We denote natural coordinates for $TM \times T^*M$ by $((x,v),(y,p))$. If we define an $\mathbb{R}^n$-valued function $f$ in these coordinates by $f((x,v),(y,p)) = (y - x)$, then $TM \oplus T^*M$ is locally defined by $f^{-1}(0)$. Thus the result will follow if we can show that $X^T \times X^{T^*}$ is in the kernel of $T_{((x,v),(y,p))}f$ for each $((x,v),(y,p)) \in f^{-1}(0)$. We compute

$$T_{((x,v),(y,p))}f((e_1,e_2),(e_3,\alpha)) = e_3 - e_1.$$

From this computation, and our local coordinate expressions for $X^T$ and $X^{T^*}$, the result follows. ■

In this way the restriction of $X^T \times X^{T^*}$ to $TM \oplus T^*M$ makes sense, and we denote the restricted LIC$^\infty$ vector field by $X^T \oplus X^{T^*}$. The following result gives the desired joint property of $X^T$ and $X^{T^*}$.

**3.4 Proposition:** *If $X$ is a LIC$^\infty$ vector field on $M$ then $X^T \oplus X^{T^*}$ leaves invariant the function $v_x \oplus \alpha_x \mapsto \alpha_x \cdot v_x$ on $TM \oplus T^*M$.*

Proof: We employ a lemma.

**1 Lemma:** *If $\tau$ is a $(1,1)$ tensor field on $M$, then the Lie derivative of the function $f_\tau\colon v_x \oplus \alpha_x \mapsto \tau(\alpha_x, v_x)$ on $TM \oplus T^*M$ with respect to the vector field $X^T \oplus X^{T^*}$ is the function $v_x \oplus \alpha_x \mapsto (\mathscr{L}_X\tau)(\alpha_x, v_x)$.*

Proof: We work in local coordinates where $f_\tau = \tau^i_j p_i v^j$. We then compute

$$\mathscr{L}_{X^T \oplus X^{T^*}} f_\tau = \frac{\partial \tau^i_j}{\partial x^k} X^k p_i v^j + \frac{\partial X^k}{\partial x^j} \tau^i_k p_i v^j - \frac{\partial X^i}{\partial x^k} \tau^k_j p_i v^j$$

which we readily verify agrees with the coordinate expression for $(\mathscr{L}_X\tau)(\alpha_x, v_x)$. ▼

We now observe that the function $v_x \otimes \alpha_x \mapsto \alpha_x \cdot v_x$ is exactly $f_{\mathrm{id}_{TM}}$ in the notation of the lemma. It thus suffices to show that $\mathscr{L}_X \mathrm{id}_{TM} = 0$ for any vector field $X$. But if we Lie differentiate the equality $\mathrm{id}_{TM}(Y) = Y$ ($Y \in \Gamma^\infty(TM)$) with respect to $X$ we obtain

$$(\mathscr{L}_X \mathrm{id}_{TM})(Y) + \mathrm{id}_{TM}([X,Y]) = [X,Y]$$

from which the proposition follows. ■

**3.5 Remark:** One may verify, in fact, that $X^{T^*}$ is the unique linear vector field on $T^*M$ which projects to $X$ and which satisfies Proposition 3.4. •

**3.4. The cotangent lift of the vertical lift.** As mentioned in the introduction, we shall deal with systems whose state space is a tangent bundle, and whose control vector fields are vertical lifts. As a consequence of an application of the maximum principle to such systems, we will be interested in the cotangent lift of vertically lifted vector fields. So let $Q$ be a finite-dimensional manifold, and let $X$ be a $\mathrm{LIC}^\infty$ vector field on $Q$ with $\mathrm{vlft}(X)$ its vertical lift to $TQ$. One computes the local coordinate expression for $\mathrm{vlft}(X)^{T^*}$ to be

$$(\mathrm{vlft}(X))^{T^*} = X^i \frac{\partial}{\partial v^i} - \frac{\partial X^j}{\partial q^i} \beta_j \frac{\partial}{\partial \alpha_i}. \tag{3.3}$$

Here we write natural coordinates for $T^*TQ$ as $((q,v),(\alpha,\beta))$.

**3.6 Remark:** It is interesting to note the relationship between $(\mathrm{vlft}(X))^{T^*}$ and $\mathrm{vlft}(X^{T^*})$. The latter vector field has the coordinate expression

$$\mathrm{vlft}(X^{T^*}) = X^i \frac{\partial}{\partial u^i} - \frac{\partial X^j}{\partial q^i} p_j \frac{\partial}{\partial \gamma_i}$$

where are writing natural coordinates for $TT^*Q$ as $((q,p),(u,\gamma))$. Now we note that there is a canonical diffeomorphism $\phi_Q$ between $T^*TQ$ and $TT^*Q$ defined in coordinates by

$$((q,v),(\alpha,\beta)) \mapsto ((q,\beta),(v,\alpha)).$$

One easily verifies that $\mathrm{vlft}(X^{T^*}) = \phi_Q^*(\mathrm{vlft}(X))^{T^*}$. We also remark that $T^*TQ$ is a symplectic manifold, since it is a cotangent bundle. Tulczyjew [1977] demonstrates that the tangent bundle of a symplectic manifold is also a symplectic manifold. Thus, in particular,

$TT^*Q$ is a symplectic manifold. The symplectic structure on $TT^*Q$ as defined by Tulczyjew is given in coordinates by

$$\omega_{TT^*Q} = \mathrm{d}q^i \wedge \mathrm{d}\gamma_i + \mathrm{d}u^i \wedge \mathrm{d}p_i.$$

One then verifies that the diffeomorphism $\phi_Q$ is symplectic with respect to these symplectic structures. Since $(\mathrm{vlft}(X))^{T^*}$ is a Hamiltonian vector field on $T^*TQ$ by Remark 3.2–1, the vector field $\mathrm{vlft}(X^{T^*})$ must also be Hamiltonian on $TT^*Q$ with the symplectic structure just described. The Hamiltonian, one readily computes, is given by $V_{\alpha_q} \mapsto \langle \alpha_q; X(q)\rangle$, where $V_{\alpha_q} \in TT^*Q$.

An intrinsic definition of $\phi_Q$ is as follows.[2] We define a map $\rho\colon T^*TQ \to T^*Q$ as follows:

$$\langle \rho(\alpha_{v_q}); u_q\rangle = \langle \alpha_{v_q}; \mathrm{vlft}_{v_q}(u_q)\rangle.$$

We may then readily verify that $\phi_Q$ is the unique map which makes the diagram

$$\begin{tikzcd}
T^*Q & TT^*Q \arrow[l, "\pi_{TT^*Q}"'] \arrow[r, "T\pi_{T^*Q}"] & TQ \\
& T^*TQ \arrow[ul, "\rho"] \arrow[u, "\phi_Q"] \arrow[ur, "\pi_{T^*TQ}"'] &
\end{tikzcd}$$

commute. •

**3.5. The canonical involution of $TTQ$.** Let $\rho_1$ and $\rho_2$ be $C^2$ maps from a neighbourhood of $(0,0) \in \mathbb{R}^2$ to $Q$. Let us denote by $(t_1, t_2)$ coordinates for $\mathbb{R}^2$. We say two such maps are ***equivalent*** if $\rho_1(0,0) = \rho_2(0,0)$ and if

$$\frac{\partial \rho_1}{\partial t_1}(0,0) = \frac{\partial \rho_2}{\partial t_1}(0,0), \qquad \frac{\partial \rho_1}{\partial t_2}(0,0) = \frac{\partial \rho_2}{\partial t_2}(0,0),$$
$$\frac{\partial^2 \rho_1}{\partial t_1 \partial t_2}(0,0) = \frac{\partial^2 \rho_2}{\partial t_1 \partial t_2}(0,0).$$

To an equivalence class $[\rho]$ we associate, in coordinates, the point

$$\left(\rho(0,0), \frac{\partial \rho}{\partial t_1}(0,0), \frac{\partial \rho}{\partial t_2}(0,0), \frac{\partial^2 \rho}{\partial t_1 \partial t_2}(0,0)\right)$$

in the local model of $TTQ$. A straightforward computation shows that this is independent of a choice of coordinates and so establishes a bijection between the set of equivalence classes and $TTQ$.

Associated with this representation of points in $TTQ$ is a canonical involution of $I_Q\colon TTQ \to TTQ$. We define $I_Q$ by saying how it acts on equivalence classes as given above. If $\rho$ is a map from a neighbourhood of $(0,0) \in \mathbb{R}^2$ to $Q$ we define $\bar{\rho}(t_1, t_2) = \rho(t_2, t_1)$ which is also then a map from a neighbourhood of $(0,0) \in \mathbb{R}^2$ into $Q$. We then define

$$I_Q([\rho]) = [\bar{\rho}].$$

In coordinates,

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

[2]The author thanks Jerry Marsden for providing this definition.

**3.6. The canonical almost tangent structure.** The final bit of tangent bundle geometry we discuss is the canonical almost tangent structure on $TM$. This is the $(1,1)$ tensor field $J_M$ on $TM$ defined by

$$J_M(X_{v_x}) = \mathrm{vlft}_{v_x}(T_{v_x}\pi_{TM}(X_{v_x}))$$

where $X_{v_x} \in T_{v_x}TM$. One verifies that in natural coordinates we have

$$J_M = \frac{\partial}{\partial v^i} \otimes \mathrm{d}x^i.$$

## 4. Some affine differential geometry

The control problems we consider are those whose drift vector field is the geodesic spray associated with an affine connection. The geometry of the affine connection enters into the optimal control problem in an interesting way, and to appreciate fully the relationship, we need a little background in affine differential geometry. We shall provide only the barest outline, and refer to Kobayashi and Nomizu [1963a, 1963b] for details.

**4.1. Basic definitions.** 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

AC1. the map $(X,Y) \mapsto \nabla_X Y$ is $\mathbb{R}$-bilinear,

AC2. $\nabla_{fX} Y = f\nabla_X Y$ for $f \in C^\infty(Q)$, and

AC3. $\nabla_X fY = f\nabla_X Y + (\mathscr{L}_X f)Y$ for $f \in C^\infty(Q)$.

The vector field $\nabla_X Y$ is called the ***covariant derivative*** of $Y$ with respect to $X$. If $(U,\phi)$ is a chart for $Q$ with coordinates $(q^1,\dots,q^n)$, then we may define $n^3$ functions $\Gamma^i_{jk}$, $i,j,k = 1,\dots,n$, on $U$ by

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

These functions are called the ***Christoffel symbols*** for the affine connection $\nabla$.

For any vector field $X$ on $Q$, one can extend $\nabla_X$ to a derivation on the entire tensor algebra on $Q$. One does this by defining $\nabla_X f = \mathscr{L}_X f$ for functions $f$, and then extending in the usual manner [Abraham, Marsden, and Ratiu 1988, Theorem 5.3.2]. Since $\nabla_X Y$ is $C^\infty(Q)$-linear in $X$, given an $(r,s)$ tensor field $\tau$ on $Q$, we may define an $(r,s+1)$ tensor field $\nabla\tau$ on $Q$ by

$$\nabla\tau(\alpha^1,\dots,\alpha^r,X_1,\dots,X_s;X_{s+1}) = (\nabla_{X_{s+1}}\tau)(\alpha^1,\dots,\alpha^r,X_1,\dots,X_s).$$

If $\tau$ is an $(r,s)$ tensor field, and $\alpha^1,\dots,\alpha^r$ are one-forms and $X_1,\dots,X_s$ are vector fields, then $\nabla\tau(\alpha^1,\dots,\alpha^r,X_1,\dots,X_s)$ denotes the one-form defined by

$$\langle \nabla\tau(\alpha^1,\dots,\alpha^r,X_1,\dots,X_s);Y\rangle = \nabla\tau(\alpha^1,\dots,\alpha^r,X_1,\dots,X_s;Y), \quad Y \in \Gamma^\infty(TQ).$$

Note that we use the semicolon prior to the last argument when we wish to employ $\nabla\tau$ as a tensor field of type $(r,s+1)$ rather than as a $T^*Q$-valued tensor of type $(r,s)$. If $X$ is a

vector field, $\nabla X$ is a $(1,1)$ tensor field which may be regarded as an endomorphism of $TQ$. We denote the dual endomorphism by $(\nabla X)^*$.

Given a Riemannian metric $g$ on $Q$, there exists a unique affine connection $\overset{g}{\nabla}$ on $Q$ having the following two properties:

LC1. $\overset{g}{\nabla} g = 0$;

LC2. $\overset{g}{\nabla}_X Y - \overset{g}{\nabla}_Y X = [X, Y]$.

The affine connection $\overset{g}{\nabla}$ is called the ***Levi-Civita connection*** associated with the Riemannian metric $g$. If $g_{ij}$ denotes the components of the Riemannian metric $g$ in a set of coordinates, then the Christoffel symbols of $\overset{g}{\nabla}$ may be verified to be

$$\overset{g}{\Gamma}{}^i_{jk} = \frac{1}{2} g^{i\ell} \left( \frac{\partial g_{\ell j}}{\partial q^k} + \frac{\partial g_{\ell k}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^\ell} \right).$$

The notion of an affine connection as we have defined it may be used to differentiate vector fields along curves. Let $c\colon I \to Q$ be a LAC curve and let $V\colon I \to TQ$ be a LAC vector field along $c$. Thus $V(t) \in T_{c(t)}Q$ for $t \in I$. If one defines a vector field $X$ which has $c$ as an integral curve, and a vector field $Y$ with the property that $Y(c(t)) = V(t)$ for $t \in I$, then we define the ***covariant derivative of V along c*** to be the vector field along $c$ defined by

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

Using the properties of affine connections, and particularly AC2, one verifies that this definition is independent of the extensions $X$ and $Y$ of $c'$ and $V$. In coordinates we have

$$(\nabla_{c'(t)} V(t))^i = \dot{V}^i(t) + \Gamma^i_{jk}(q(t)) \dot{q}^j(t) V^k(t).$$

In like manner we may define the covariant derivative of a general tensor field along $c$.

Of particular interest are those LAC curves $c\colon I \to Q$ for which $\nabla_{c'(t)} c'(t) = 0$. Such curves are called ***geodesics*** for the affine connection $\nabla$. Thus a geodesic in coordinates satisfies the equation

$$\ddot{q}^i(t) + \Gamma^i_{jk}(q(t)) \dot{q}^j(t) \dot{q}^k(t) = 0, \quad i = 1, \dots, n.$$

Geodesics are thus curves of class $C^\infty$. Note that the geodesic equation is a second-order differential equation, which therefore defines a first-order vector field on $TQ$. We denote this first-order vector field by $Z$ and note that in coordinates

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

This vector field is called the ***geodesic spray*** of $\nabla$. The projections of integral curves of $Z$ to $Q$ by the tangent bundle projection are exactly the geodesics of $\nabla$.

**4.2. Torsion and curvature.** The condition LC2 suggests that for a general affine connection the object

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

may be of some importance. Indeed it is, and $T$ so defined is a $(1,2)$ tensor field called the ***torsion*** for $\nabla$.

We shall also be interested in another tensor which one may associate to an affine connection: the ***curvature tensor***. We shall attach some geometric significance to the curvature tensor shortly, so for now we simply give its definition. The curvature tensor is a $(1,3)$ tensor field on $Q$, but it is convenient to give it somewhat non-standard notation. If the $(1,3)$ tensor field is $\tilde{R}$, then we may define a section $R$ of $\mathrm{T}^0_2(TQ) \otimes \mathrm{T}^1_1(TQ)$ by $R(X,Y)W = \tilde{R}(X,Y,W)$. With this notation we define $R$ by

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

One verifies that $R$ satisfies the ***first Bianchi identity***:

$$\sum_{\sigma \in S_3} \big(R(X_{\sigma(1)}, X_{\sigma(2)}) X_{\sigma(3)}\big) = \sum_{\sigma \in S_3} \big(T(T(X_{\sigma(1)}, X_{\sigma(2)}), X_{\sigma(3)}) + (\nabla_{X_{\sigma(1)}} T)(X_{\sigma(2)}, X_{\sigma(3)})\big) \tag{4.1}$$

where $S_3$ is the permutation group on three symbols.

It will be useful for us to have coordinate expressions for the torsion and curvature tensors:

$$\begin{aligned} T^i_{jk} &= \Gamma^i_{jk} - \Gamma^i_{kj} \\ R^i_{jk\ell} &= \frac{\partial \Gamma^i_{\ell j}}{\partial q^k} - \frac{\partial \Gamma^i_{kj}}{\partial q^\ell} + \Gamma^i_{km}\Gamma^m_{\ell j} - \Gamma^i_{\ell m}\Gamma^m_{kj}. \end{aligned} \tag{4.2}$$

In writing these coordinate formulas we use the conventions

$$T(X,Y) = T^i_{jk} X^j Y^k, \quad R(X,Y)W = R^i_{jk\ell} W^j X^k Y^\ell.$$

We will also need the coordinate formula for $\nabla T$:

$$(\nabla T)^i_{jkl} = \frac{\partial \Gamma^i_{jk}}{\partial q^\ell} - \frac{\partial \Gamma^i_{kj}}{\partial q^\ell} + \Gamma^i_{\ell m}\Gamma^m_{jk} - \Gamma^i_{\ell m}\Gamma^m_{kj} - \Gamma^i_{mk}\Gamma^m_{\ell j} + \Gamma^i_{km}\Gamma^m_{\ell j} - \Gamma^i_{jm}\Gamma^m_{\ell k} + \Gamma^i_{mj}\Gamma^m_{\ell k}. \tag{4.3}$$

This formula uses the convention

$$\nabla T(W,X;Y) = (\nabla T)^i_{jk\ell} W^j X^k Y^\ell,$$

which, we remark, differs from our convention used in writing the components for $R$.

It will be helpful to distinguish when we are talking about the curvature tensor for the Levi-Civita connection. Thus we denote this by $\overset{g}{R}$. The curvature tensor for the Levi-Civita connection has some symmetries which will be useful for us. In particular, the following formula may be found in [Kobayashi and Nomizu 1963a, Chapter V, Proposition 2.1]:

$$g(\overset{g}{R}(X_3,X_4)X_2, X_1) = g(\overset{g}{R}(X_1,X_2)X_4, X_3) \tag{4.4}$$

for all vector fields $X_1, \dots, X_4 \in \Gamma^\infty(TQ)$.

**4.3. The Jacobi equation.** We will be dealing with control systems, affine in the controls, for which the drift vector field is the geodesic spray of an affine connection. When studying the optimal control problem for such systems, one is led to equations which are "dual" to the equations describing geodesic deviation. Here we consider the equations of geodesic variation themselves. Thus one considers a situation as follows. Let $\nabla$ be an affine connection on $Q$ and fix a geodesic $c\colon I \to Q$ for $\nabla$. By a ***differentiable family of geodesics around c*** we mean a map $\sigma\colon I \times [-\epsilon, \epsilon] \to Q$ with the following properties:

1. $\sigma$ is differentiable;
2. $t \mapsto \sigma(t, s)$ is a geodesic for $\nabla$ for every $s \in [-\epsilon, \epsilon]$;
3. $\sigma(t, 0) = c(t)$ for every $t \in I$.

A ***Jacobi field*** is then a vector field $\xi$ along a geodesic $c$ of the form

$$\xi(t) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \sigma(t, s), \quad t \in I$$

for some differentiable family $\sigma$ of geodesics around $c$. One verifies that a necessary and sufficient condition for a vector field $\xi$ along a geodesic to be a Jacobi field is that it satisfy the ***Jacobi equation***:

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

Note that this equation is a second-order equation for the vector field $\xi$ along $c\colon I \to Q$. Thus to specify its solution uniquely requires an initial condition for $\xi(a)$ and $\nabla_{c'(a)}\xi(a)$ for some $a \in I$.

## 5. Ehresmann connections induced by an affine connection

This section concludes the first part of the paper, and provides the essential ingredient of the maximum principle for affine connection control systems: the adjoint Jacobi equation. In actuality, the results of this section, particularly those of Section 5.9, represent the meat of the paper since the optimal control results of Part II follow in a fairly straightforward way once one has at one's disposal the results which we now provide.

**5.1. Motivating remarks.** As we have stated several times already, we will be looking at control affine systems whose drift vector field is the geodesic spray $Z$ for an affine connection. Readers familiar with the geometry of the maximum principle will immediately realise that the cotangent lift of $Z$ will be important for us. One way to frame the objective of this section is to think about how one might represent $Z^{T^*}$ in terms of objects defined on $Q$, even though $Z^{T^*}$ is itself a vector field on $T^*TQ$. That this ought to be possible seems reasonable as all the information used to describe $Z^{T^*}$ is contained in the affine connection $\nabla$ on $Q$, along with some canonical tangent and cotangent bundle geometry. It turns out that it *is* possible to essentially represent $Z^{T^*}$ on $Q$, but to do so requires some effort. What's more, it is perhaps not immediately obvious how one should proceed. A possible way to get moving in the right direction is as follows:

1. according to Remark 3.1–4 and the very definition of the Jacobi equation, we expect there to be some relationship between $Z^T$ and the Jacobi equation;

2. by Remark 3.5 there is a relationship between $Z^T$ and $Z^{T^*}$;

3. from 1 and 2 we may expect that by coming to understand the relationship between $Z^T$ and the Jacobi equation, one may be able to see how to essentially represent $Z^{T^*}$ on $Q$.

One sees, then, that our approach to understanding $Z^{T^*}$ entails that we first understand $Z^T$ in terms of the Jacobi equation. We shall see that once we have done this, it is a simple matter to "dualise" our constructions to arrive at what we shall call the adjoint Jacobi equation, a one-form version of the Jacobi equation which appears in the maximum principle for affine connection control systems.

In order to simultaneously understand $Z^T$ and $Z^{T^*}$ we use various Ehresmann connections to provide splittings of the necessary tangent spaces. Note that as maps the vector fields $Z^T$ and $Z^{T^*}$ are $T(TTQ)$ and $T(T^*TQ)$-valued, respectively. Also note that the tangent spaces to $TTQ$ and $T^*TQ$ are $4\dim(Q)$-dimensional. Our goal is to break these tangent spaces up into four parts, each of dimension $\dim(Q)$. Some of the Ehresmann connections we describe here are well-known, but others might be new, even though they are straightforward to describe. We refer to [Kolář, Michor, and Slovák 1993, Chapter III] for a general discussion of Ehresmann connections. Along the way we will point out various interesting relationships between the objects we encounter. Some of these relationships are revealed in the book of Yano and Ishihara [1973].

The reader wishing to cut to the chase and see the point of producing all of these Ehresmann connections is referred forward to Theorems 5.6 and 5.10.

**5.2. Ehresmann connections.** The reader will recall that an ***Ehresmann connection*** on a locally trivial fibre bundle $\pi\colon M \to B$ is a complement $HM$ to $VM \triangleq \ker(T\pi)$ in $TM$. Thus $TM = HM \oplus VM$. We denote by $\mathrm{hor}\colon TM \to TM$ the horizontal projection, and by $\mathrm{ver}\colon TM \to TM$ the vertical projection. Note that for each $x \in M$, $T_x\pi|H_xM\colon H_xM \to T_{\pi(x)}B$ is an isomorphism. We denote its inverse by $\mathrm{hlft}_x\colon T_{\pi(x)}B \to H_xM$. If $((x^1,\dots,x^m),(y^1,\dots,y^{n-m}))$ are bundle coordinates for $M$ then we have

$$\mathrm{hlft}_{(x,y)}\Big(\frac{\partial}{\partial x^a}\Big) = \frac{\partial}{\partial x^a} + C_a^\alpha(x,y)\frac{\partial}{\partial y^\alpha}, \quad a = 1,\dots,m.$$

This defines the connection coefficients $C_a^\alpha$, $\alpha = 1,\dots,n-m$, $a = 1,\dots,m$.

Given an Ehresmann connection, the ***connection form*** is the vertical-valued one-form $\omega$ defined by

$$\omega(v_x) = \mathrm{ver}(v_x).$$

The ***curvature form*** is the vertical-valued two-form $\Omega$ given by

$$\Omega(u_x, v_x) = -\omega([X,Y](x))$$

where $X$ and $Y$ are vector fields which extend $u_x$ and $v_x$, respectively. One verifies that $\Omega$ does not depend on the extensions, and so only depends on the value of $X$ and $Y$ at $x \in M$.

**5.3. Linear connections on vector bundles.** In this section we generalise the constructions of Section 3.3.

If $\pi\colon E \to B$ is a vector bundle, then $VE$ is isomorphic to the pull-back bundle $\pi^*\pi\colon \pi^*E \to E$ whose fibre over $e_b \in E$ is exactly $\pi^{-1}(b)$. Thus the projection ver associated with an Ehresmann connection $HE$ may be thought of as a map $\mathrm{ver}\colon TE \to E$. Further, we define a vector bundle isomorphism $\mathrm{vlft}\colon \pi^*E \to VE$ by

$$\mathrm{vlft}(\tilde{e}_b, e_b) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\tilde{e}_b + te_b).$$

We adopt the notation of Section 3.1 and write $\mathrm{vlft}_{\tilde{e}_b}(e_b)$ rather then $\mathrm{vlft}(\tilde{e}_b, e_b)$. An Ehresmann connection on $\pi\colon E \to B$ is called ***linear*** if $\mathrm{ver}\colon TE \to E$ is a vector bundle mapping with respect to the vector bundles $T\pi\colon TE \to TB$ and $\pi_{TE}\colon TE \to E$, and if $\mathrm{ver} \circ \mathrm{vlft} = \pi_{TE}$ (where we think of ver as being $E$-valued). One verifies that an Ehresmann connection is linear if and only if the connection coefficients have the form $C^\alpha_a(x,u) = A^\alpha_{a\beta}(x)u^\beta$ where $(x,u)$ are vector bundle coordinates. This defines local functions $A^\alpha_{a\beta}$, $\alpha, \beta = 1, \dots, n-m$, $a = 1, \dots, m$, defined on the base space.

Let $X$ be a vector field on the base space $B$ of a vector bundle $\pi\colon E \to B$. A ***linear vector field*** over $X$ is a vector field $Y\colon E \to TE$ which is $\pi$-related to $X$ and which is a vector bundle mapping. In coordinates $(x,u)$ a linear vector field $Y$ over $X$ has the form

$$Y = X^a(x)\frac{\partial}{\partial x^a} + Y^\alpha_\beta(x)u^\beta\frac{\partial}{\partial u^\alpha} \tag{5.1}$$

where $X^a$ are the components of $X$ and for some functions $Y^\alpha_\beta$, $\alpha, \beta = 1, \dots, n-m$. The flow of a linear vector field is comprised of local vector bundle isomorphisms of $\pi\colon E \to M$. Thus, if $\pi^*\colon E^* \to B$ is the dual bundle and if $\nu$ is a linear vector field on $E^*$ over $X$, we may define $Y \oplus \nu$ as the linear vector field on $E \oplus E^*$ whose flow is the family of vector bundle isomorphisms given by the direct sum of those generated by $Y$ and by $\nu$. We may define a function $f_E$ on $E \oplus E^*$ by $f_E(e_b \oplus \alpha_b) = \alpha_b \cdot e_b$. If $Y$ is a linear vector field over $X$ on $E$ then there exists a unique linear vector field over $X$ on $E^*$, denoted $Y^*$, with the property that

$$\mathscr{L}_{Y \oplus Y^*} f_E = 0.$$

In coordinates, if $Y$ is as given by (5.1) then

$$Y^* = X^a(x)\frac{\partial}{\partial x^a} - Y^\alpha_\beta(x)\rho_\alpha\frac{\partial}{\partial \rho_\beta}.$$

Of course, when $E = TM$ and $Y = X^T$, we see that $Y^* = X^{T^*}$, consistent with Remark 3.5.

If $HE$ is a linear connection on a vector bundle $\pi\colon E \to B$ then there exists a unique connection $HE^*$ on the dual bundle $\pi^*\colon E^* \to B$ which satisfies the property

$$\mathrm{hlft}(X)^* = \mathrm{hlft}^*(X) \tag{5.2}$$

where $X$ is a vector field on $B$, hlft is the horizontal lift associated with $HE$, and $\mathrm{hlft}^*$ is the horizontal lift associated with $HE^*$ [Kolář, Michor, and Slovák 1993, §47.15]. If $A^\alpha_{a\beta}(x)u^\beta$ are the connection coefficients for $HE$ in vector bundle coordinates $(x,u)$, then the connection coefficients for $HE^*$ in the dual vector bundle coordinates $(x,\rho)$ are $-A(x)^\beta_{a\alpha}\rho_\beta$.

**5.4. The Ehresmann connection on $\pi_{TM}\colon TM \to M$ associated with a second-order vector field on $TM$.** Let $S$ be a second-order vector field on $TM$. Thus $S\colon TM \to TTM$ has the property that $T\pi_{TM} \circ X = \mathrm{id}_{TM}$. We define an Ehresmann connection on $\pi_{TM}\colon TM \to M$ as follows [Crampin 1983]. Recall from Section 3.6 the canonical almost tangent structure $J_M$ on $TM$. One may verify that the kernel of the endomorphism $\mathscr{L}_S J_M\colon TTM \to TTM$ is a subbundle complementary to $VTM = \ker(T\pi_{TM})$. We denote this complementary distribution, which is thus an Ehresmann connection, by $HTM$. In natural coordinates $(x, v)$ for $TM$, we may write a second order vector field $S$ as

$$S = v^i \frac{\partial}{\partial x^i} + S^i(x, v) \frac{\partial}{\partial v^i}.$$

One then verifies that a local basis for $HTM$ is given by the vector fields

$$\mathrm{hlft}\Big(\frac{\partial}{\partial x^i}\Big) = \frac{\partial}{\partial x^i} + \frac{1}{2}\frac{\partial S^j}{\partial v^i}\frac{\partial}{\partial v^j}, \quad i = 1, \dots, n. \tag{5.3}$$

This Ehresmann connection gives a splitting of $T_{v_x}TM$ into a horizontal and a vertical part. The horizontal part is isomorphic to $T_xM$ via $\mathrm{hlft}_{v_x}$, and the vertical part is isomorphic to $T_xM$ in the natural way (it is the tangent space to a vector space). Thus we have a natural isomorphism $T_{v_x}TM \simeq T_xM \oplus T_xM$, and we adopt the convention that the first part of this splitting will be horizontal, and the second will be vertical.

**5.5. The Ehresmann connection on $\pi_{TQ}\colon TQ \to Q$ associated with an affine connection on $Q$.** If $Z$ is the geodesic spray defined by an affine connection $\nabla$ on $Q$, then we may use the construction of the previous section to provide an Ehresmann connection on $\pi_{TQ}\colon TQ \to Q$. One further verifies that in local coordinates, a basis for $HTQ$ is given by

$$\mathrm{hlft}\Big(\frac{\partial}{\partial q^i}\Big) = \frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})v^k \frac{\partial}{\partial v^j}, \quad i = 1, \dots, n.$$

This defines "hlft" as the horizontal lift map for the connection we describe here. We shall introduce different notation for the horizontal lift associated with the other Ehresmann connections we define. We also denote by "vlft" the vertical lift map associated with this connection. Note that we have $Z(v_q) = \mathrm{hlft}_{v_q}(v_q)$.

The Ehresmann connection $HTQ$ defines a connection form $\omega$ and a curvature form $\Omega$, just as in Section 5.2. It will be useful to have a formula relating $\Omega$ to the curvature tensor $R$ and the torsion tensor $T$.

**5.1 Proposition:** *Let $\nabla$ be an affine connection on $Q$ and let $\Omega$ be the curvature form for the associated Ehresmann connection on $\pi_{TQ}\colon TQ \to Q$. The following formula holds:*

$$\begin{aligned}\Omega(\mathrm{hlft}_{v_q}(u_q), \mathrm{hlft}_{v_q}(w_q)) = \mathrm{vlft}_{v_q}\big(&R(u_q, w_q)v_q - \tfrac{1}{2}(\nabla_{u_q}T)(w_q, v_q) + \\ &\tfrac{1}{2}(\nabla_{w_q}T)(u_q, v_q) - \tfrac{1}{2}T(T(u_q, w_q), v_q) + \tfrac{1}{4}(T(T(u_q, v_q), w_q) - T(T(w_q, v_q), u_q)\big).\end{aligned}$$

*In particular, if $\nabla$ is torsion-free then*

$$\Omega(\mathrm{hlft}_{v_q}(u_q), \mathrm{hlft}_{v_q}(w_q)) = \mathrm{vlft}_{v_q}\big(R(u_q, w_q)v_q\big).$$

Proof: The most straightforward, albeit tedious, proof is in coordinates. Let $X$ and $Y$ be vector fields which extend $u_q$ and $w_q$, respectively. A computation yields

$$\begin{aligned}\operatorname{ver}([\operatorname{hlft} X, \operatorname{hlft}(Y)]) = \frac{1}{2}\Big(&\frac{\partial \Gamma^i_{jk}}{\partial q^\ell} + \frac{\partial \Gamma^i_{kj}}{\partial q^\ell} - \frac{\partial \Gamma^i_{j\ell}}{\partial q^k} - \frac{\partial \Gamma^i_{\ell j}}{\partial q^k} + \\ &\frac{1}{2}\Big(\Gamma^i_{m\ell}\Gamma^m_{kj} + \Gamma^i_{m\ell}\Gamma^m_{jk} + \Gamma^i_{\ell m}\Gamma^m_{kj} + \Gamma^i_{\ell m}\Gamma^m_{jk} - \\ &\Gamma^i_{mk}\Gamma^m_{\ell j} - \Gamma^i_{mk}\Gamma^m_{j\ell} - \Gamma^i_{km}\Gamma^m_{\ell j} - \Gamma^i_{km}\Gamma^m_{j\ell}\Big)\Big) v^j X^k Y^l \frac{\partial}{\partial v^i}.\end{aligned}$$

One now employs the coordinate formulas (4.2) for $T$ and $R$, and the coordinate formula (4.3) for $\nabla T$ to directly verify that

$$\begin{aligned}\operatorname{ver}([\operatorname{hlft} X, \operatorname{hlft}(Y)])(v_q) = \operatorname{vlft}_{v_q}\big(&R(w_q, u_q)v_q + R(w_q, v_q)u_q + R(v_q, u_q)w_q + \\ &(\nabla_{v_q} T)(w_q, u_q) + \tfrac{1}{2}T(T(u_q, v_q), w_q) + \tfrac{1}{2}T(T(v_q, w_q), u_q).\end{aligned}$$

The result may now be proved using the first Bianchi identity (4.1). ■

Recall that the connection on $\pi_{TQ}\colon TQ \to Q$ provides a natural isomorphism of $T_{v_q}TQ$ with $T_qQ \oplus T_qQ$ as in Section 5.4. This in turn provides an isomorphism of $T^*_{v_q}TQ$ with $T^*_qQ \oplus T^*_qQ$. In coordinates the basis

$$\mathrm{d}q^i, \quad \mathrm{d}v^i + \frac{1}{2}(\Gamma^i_{jk} + \Gamma^i_{kj})v^k \mathrm{d}q^j, \quad i = 1, \dots, n, \tag{5.4}$$

is adapted to this splitting in that the first $n$ vectors form a basis for the horizontal part of $T^*_{v_q}TQ$ and the second $n$ vectors form a basis for the vertical part.

**5.6. The Ehresmann connection on $\pi_{T^*Q}\colon T^*Q \to Q$ associated with an affine connection on $Q$.** We can equip the cotangent bundle of $Q$ with an Ehresmann connection induced by $\nabla$. We do this by noting that the Ehresmann connection $HTQ$ induced by an affine connection is a linear connection. Thus, we have an induced connection $HT^*Q$ on $\pi_{T^*Q}\colon T^*Q \to Q$ whose horizontal lift is as defined by (5.2). A coordinate basis for the $HT^*Q$ is given by

$$\operatorname{hlft}^*\Big(\frac{\partial}{\partial q^i}\Big) = \frac{\partial}{\partial q^i} + \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})p_j \frac{\partial}{\partial p_k}, \quad i = 1, \dots, n.$$

Note that $\operatorname{hlft}^*$ denotes the horizontal lift for the Ehresmann connection on $\pi_{T^*Q}\colon T^*Q \to Q$. The vertical subbundle has the basis

$$\operatorname{vlft}^*\big(\mathrm{d}q^i\big) = \frac{\partial}{\partial p_i}, \quad i = 1, \dots, n$$

which, we remark, defines the vertical lift map $\operatorname{vlft}^*$.

**5.7. The Ehresmann connection on $\pi_{TTQ}\colon TTQ \to TQ$ associated with an affine connection on $Q$.** In this section we induce a connection on $\pi_{TTQ}\colon TTQ \to TQ$ using a second-order vector field as described in Section 5.4. We do not yet *have* a second-order vector field on $TTQ$, so let us set about producing one from the geodesic spray $Z$. Note that $Z^T$ is a vector field on $TTQ$, and we may determine its coordinate expression to be

$$Z^T = v^i \frac{\partial}{\partial q^i} - \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i} + w^i \frac{\partial}{\partial u^i} - \left( \frac{\partial \Gamma^i_{jk}}{\partial q^\ell} v^j v^k u^\ell + \Gamma^i_{jk} w^j v^k + \Gamma^i_{kj} w^j v^k \right) \frac{\partial}{\partial w^i}. \tag{5.5}$$

This vector field is "almost" second-order. To make it second-order, we use the canonical involution $I_Q\colon TTQ \to TTQ$ described in Section 3.5 and the following lemma.

**5.2 Lemma:** *$I_Q^* Z^T$ is a second-order vector field on $TTQ$.*

Proof: A simple computation, using the coordinate expressions for $I_Q$ and $Z^T$, gives

$$I_Q^* Z^T = u^i \frac{\partial}{\partial q^i} + w^i \frac{\partial}{\partial v^i} - \Gamma^i_{jk} u^j u^k \frac{\partial}{\partial u^i} - \left( \frac{\partial \Gamma^i_{jk}}{\partial q^\ell} v^\ell u^j u^k + \Gamma^i_{jk} u^k w^j + \Gamma^i_{kj} u^k w^j \right) \frac{\partial}{\partial w^i} \tag{5.6}$$

which verifies the lemma. ■

Now we may use the procedure of Section 5.4 to produce a connection on $\pi_{TTQ}\colon TTQ \to TQ$. Let us denote this connection by $H(TTQ)$. Note that this connection provides a splitting

$$T_{X_{v_q}} TTQ \simeq T_{v_q} TQ \oplus T_{v_q} TQ \tag{5.7}$$

for $X_{v_q} \in T_{v_q} TQ$. Also, the connection $HTQ$ on $\pi_{TQ}\colon TQ \to Q$ described in Section 5.5 gives a splitting $T_{v_q} TQ \simeq T_q Q \oplus T_q Q$. Therefore, we have a resulting splitting

$$T_{X_{v_q}} TTQ \simeq T_q Q \oplus T_q Q \oplus T_q Q \oplus T_q Q. \tag{5.8}$$

In this splitting, the first two components are the horizontal subspace and the second two components are the vertical subspace. Within each pair, the first part is horizontal and the second is vertical.

Let us now write a basis of vector fields on $TTQ$ which is adapted to the splitting (5.8). To obtain a coordinate expression for a basis of this connection, we use the coordinate expression (5.6) for $I_Q^* Z^T$. We write a basis which is adapted to the splitting of $T_{v_q} TQ$. The resulting basis vectors for $H(TTQ)$ are

$$\begin{aligned}
\operatorname{hlft}^T\Big( \frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki}) v^k \frac{\partial}{\partial v^j} \Big) &= \frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki}) v^k \frac{\partial}{\partial v^j} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki}) u^k \frac{\partial}{\partial u^j} - \\
&\quad \frac{1}{2}\Big( \frac{\partial \Gamma^j_{i\ell}}{\partial q^k} u^\ell v^k + \frac{\partial \Gamma^j_{\ell i}}{\partial q^k} u^\ell v^k + (\Gamma^j_{ik} + \Gamma^j_{ki}) w^k - \\
&\quad \tfrac{1}{2}(\Gamma^k_{i\ell} + \Gamma^k_{\ell i})(\Gamma^j_{km} + \Gamma^j_{mk}) u^m v^\ell \Big) \frac{\partial}{\partial w^j}, \quad i = 1, \dots, n, \\
\operatorname{hlft}^T\Big( \frac{\partial}{\partial v^i} \Big) &= \frac{\partial}{\partial v^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki}) u^k \frac{\partial}{\partial w^j}, \quad i = 1, \dots, n,
\end{aligned}$$

with the first $n$ basis vectors forming a basis for the horizontal part of $H_{X_{v_q}}(TTQ)$ and the second $n$ vectors forming a basis for the vertical part of $H_{X_{v_q}}(TTQ)$, with respect to the splitting $H_{X_{v_q}}(TTQ) \simeq T_qQ \oplus T_qQ$. Note that we use the notation $\mathrm{hlft}^T$ to refer to the horizontal lift for the connection on $\pi_{TTQ}\colon TTQ \to TQ$. We also denote by $\mathrm{vlft}^T$ the vertical lift on this vector bundle.

We may easily derive a basis for the vertical subbundle of $\pi_{TTQ}\colon TTQ \to TQ$ which adapts to the splitting of $T_{v_q}TQ$. We verify that the vector fields

$$\mathrm{vlft}^T\Big(\frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})v^k\frac{\partial}{\partial v^j}\Big) = \frac{\partial}{\partial u^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})v^k\frac{\partial}{\partial w^j}, \quad i = 1, \dots, n,$$
$$\mathrm{vlft}^T\Big(\frac{\partial}{\partial v^i}\Big) = \frac{\partial}{\partial w^i}, \quad i = 1, \dots, n,$$

have the property that the first $n$ vectors span the horizontal part of $V_{X_{v_q}}TTQ$, and the second $n$ span the vertical part of $V_{X_{v_q}}TTQ$.

**5.3 Remarks:** 1. The construction in this section may, in fact, be made with an arbitrary second-order vector field. That is, if $S$ is a second-order vector field on $TQ$, then $I_Q^*S^T$ is a second-order vector field on $TTQ$. This second-order vector field then induces a connection on $\pi_{TTQ}\colon TTQ \to TQ$. Clearly then, this construction can be iterated and so provides a connection on the $k$th tangent bundle $\pi_{TT^{k-1}Q}\colon T^kQ \to T^{k-1}Q$ for each $k \geq 0$. This then provides an isomorphism of $T_XT^{k-1}Q$ with the direct sum $T_qQ \oplus \cdots \oplus T_qQ$ of $2k$ copies of $T_qQ$ where $X \in T^{k-1}Q$ and $q = \pi_{TQ} \circ \cdots \circ \pi_{TT^{k-1}Q}(X)$.

2. The vector field $I_Q^*Z^T$ is the geodesic spray of an affine connection on $TQ$. Thus we see how, given an affine connection on a manifold $Q$, it is possible to derive an affine connection on the $k$th tangent bundle $T^kQ$ for $k \geq 1$.

3. The vector field $I_Q^*Z^T$ appears to be the geodesic spray of the affine connection on $TQ$ which Yano and Ishihara [1973] call the "complete lift" of the affine connection $\nabla$ on $Q$. This affine connection is defined as the unique affine connection $\nabla^T$ on $TQ$ which satisfies $\nabla^T_{X^T}Y^T = \big(\nabla_XY\big)^T$ for vector fields $X$ and $Y$ on $Q$. Yano and Ishihara also show that the geodesics of $\nabla^T$ are vector fields along geodesics of $\nabla$. They further claim that these vector fields along geodesics are in fact Jacobi fields. This is indeed true as our results of Section 5.9 show, but the proof in Yano and Ishihara cannot be generally correct as, for example, the Jacobi equation they use lacks the torsion term which the actual Jacobi equation possesses. •

**5.8. The Ehresmann connection on $\pi_{T^*TQ}\colon T^*TQ \to TQ$ associated with an affine connection on $Q$.** As the tangent bundle of $TQ$ comes equipped with an Ehresmann connection, so too does its cotangent bundle. To construct this connection, we note that the connection $H(TTQ)$ on $\pi_{TTQ}\colon TTQ \to TQ$ is a linear connection and thus there is a connection naturally induced on $\pi_{T^*TQ}\colon T^*TQ \to TQ$ whose horizontal lift is as defined by (5.2). We denote this connection by $H(T^*TQ)$, and note that it provides a splitting

$$T_{\Lambda_{v_q}}T^*TQ \simeq T_{v_q}TQ \oplus T^*_{v_q}TQ$$

for $\Lambda_{v_q} \in T^*TQ$. In turn, the connection of Section 5.5 on $\pi_{TQ}\colon TQ \to Q$ gives a splitting $T_{v_q}TQ \simeq T_qQ \oplus T_qQ$ and so also a splitting $T^*_{v_q}TQ \simeq T^*_qQ \oplus T^*_qQ$. This then provides the

splitting

$$T_{\Lambda_{\alpha_q}}T^*TQ \simeq T_qQ \oplus T_qQ \oplus T_q^*Q \oplus T_q^*Q.$$

The first two components of this splitting are the horizontal part of the subspace, and the second two are the vertical part. For each pair, the first component is horizontal and the second is vertical.

Let us now write a basis for vector fields on $T^*TQ$ which is adapted to the splitting we have just demonstrated. First we determine that a basis for the horizontal subbundle is

$$\begin{aligned}
\operatorname{hlft}^{T^*}\Big(\frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})v^k\frac{\partial}{\partial v^j}\Big) &= \frac{\partial}{\partial q^i} - \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})v^m\frac{\partial}{\partial v^j} + \\
&\quad \frac{1}{2}\Big(\frac{\partial\Gamma^j_{ik}}{\partial q^\ell}v^\ell\beta_j + \frac{\partial\Gamma^j_{ki}}{\partial q^\ell}v^\ell\beta_j + (\Gamma^j_{ik} + \Gamma^j_{ki})\alpha_j - \\
&\quad \frac{1}{2}(\Gamma^\ell_{im} + \Gamma^\ell_{mi})(\Gamma^j_{\ell k} + \Gamma^j_{k\ell})v^m\beta_j\Big)\frac{\partial}{\partial\alpha_k} + \\
&\quad \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})\beta_j\frac{\partial}{\partial\beta_k}, \quad i = 1,\dots,n, \\
\operatorname{hlft}^{T^*}\Big(\frac{\partial}{\partial v^i}\Big) &= \frac{\partial}{\partial v^i} + \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})\beta_j\frac{\partial}{\partial\alpha_k}, \quad i = 1,\dots,n.
\end{aligned}$$

The first $n$ basis vectors span the horizontal part of $H_{\Lambda_{v_q}}T^*TQ$, and the second $n$ vectors span the vertical part. We have introduced the notation $\operatorname{hlft}^{T^*}$ to refer to the horizontal lift map on the bundle $\pi_{T^*TQ}\colon T^*TQ \to TQ$, and we shall denote the vertical lift map by $\operatorname{vlft}^{T^*}$.

One may also write a basis for $V(T^*TQ) = \ker(T\pi_{T^*TQ})$ which is adapted to the splitting of $T^*_{v_q}TQ$. We use (5.4) to provide a basis

$$\begin{aligned}
\operatorname{vlft}^{T^*}(\mathrm{d}q^i) &= \frac{\partial}{\partial\alpha_i}, \quad i = 1,\dots,n, \\
\operatorname{vlft}^{T^*}\Big(\mathrm{d}v^i + \frac{1}{2}(\Gamma^i_{jk} + \Gamma^i_{kj})v^k\mathrm{d}q^j\Big) &= \frac{\partial}{\partial\beta_i} + \frac{1}{2}(\Gamma^i_{jk} + \Gamma^i_{kj})v^k\frac{\partial}{\partial\alpha_j}, \quad i = 1,\dots,n,
\end{aligned}$$

in which the first $n$ vectors are a basis for the horizontal part of $V_{\Lambda_{v_q}}T^*TQ$ and the second $n$ are a basis for the vertical part.

**5.9. Representations of $Z^T$ and $Z^{T^*}$.** In the previous sections we have provided local bases for the various connections we constructed. With these local bases in hand, and with the coordinate expressions (5.5) and (5.12) (see below) for $Z^T$ and $Z^{T^*}$, it is a simple matter to determine the form of $Z^T(X_{v_q})$ and $Z^{T^*}(\Lambda_{v_q})$ in these splittings, where $X_{v_q} \in TTQ$ and $\Lambda_{v_q} \in T^*TQ$. Let us merely record the results of these somewhat tedious computations.

First we look at $Z^T$. In this case recall that the connection $H(TTQ)$ on $\pi_{TTQ}\colon TTQ \to TQ$ and the connection $HTQ$ on $\pi_{TQ}\colon TQ \to Q$ combine to give a splitting

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

where $X_{v_q} \in T_{v_q}TQ$. Here we maintain our convention that the first two components refer to the horizontal component for a connection $H(TTQ)$ on $\pi_{TTQ}\colon TTQ \to TQ$, and the

second two components refer to the vertical component. Using the splitting (5.7) let us write $X_{v_q} \in T_{v_q}TQ$ as $u_{v_q} \oplus w_{v_q}$ for some $u_{v_q}, w_{v_q} \in T_qQ$. Note that we depart from our usual notation of writing tangent vectors in $T_qQ$ with a subscript of $q$, instead using the subscript $v_q$. This abuse of notation is necessary (and convenient) to reflect the fact that these vectors depend on where we are in $TQ$ and not just in $Q$. A computation verifies the following result, where $\Omega$ is the curvature form for the connection $HTQ$.

**5.4 Proposition:** $Z^T(u_{v_q} \oplus w_{v_q}) = v_q \oplus 0 \oplus w_{v_q} \oplus \big(-\Omega(\mathrm{hlft}_{v_q}(u_{v_q}), \mathrm{hlft}_{v_q}(v_q))\big)$.

In writing this formula, we are regarding $\Omega$ as taking values in $V_{v_q}TQ \simeq T_qQ$.

We may use this representation of $Z^T$ to obtain a refined relationship between solutions of the Jacobi equation and integral curves of $Z^T$. To do so, we first prove a simple lemma. We state a more general form of this lemma than we shall immediately use, but the extra generality will be useful in Section 7.

**5.5 Lemma:** *Let $Y$ be a* LIC$^\infty$ *vector field on $Q$ and suppose that $c\colon I \to Q$ is the* LAD *curve satisfying $\nabla_{c'(t)}c'(t) = Y(t, c(t))$, and denote by $\sigma\colon I \to TQ$ the tangent vector field of $c$ (i.e., $\sigma = c'$). Let $X\colon I \to TTQ$ be a* LAC *vector field along $\sigma$, and denote $X(t) = X_1(t) \oplus X_2(t) \in T_{c(t)}Q \oplus T_{c(t)}Q \simeq T_{\sigma(t)}TQ$. Then the tangent vector to the curve $t \mapsto X(t)$ is given by $c'(t) \oplus Y(t, c(t)) \oplus \tilde{X}_1(t) \oplus \tilde{X}_2(t)$ where*

$$\tilde{X}_1(t) = \nabla_{c'(t)}X_1(t) + \tfrac{1}{2}T(X_1(t), c'(t)), \qquad \tilde{X}_2(t) = \nabla_{c'(t)}X_2(t) + \tfrac{1}{2}T(X_2(t), c'(t)).$$

Proof: In coordinates the curve $t \mapsto X(t)$ has the form

$$(q^i(t), \dot{q}^j(t), X_1^k(t), X_2^\ell(t) - \tfrac{1}{2}(\Gamma^\ell_{mr} + \Gamma^\ell_{rm})\dot{q}^m(t)X_1^r(t)).$$

The tangent vector to this curve is then given a.e. by

$$\dot{q}^i\frac{\partial}{\partial q^i} + (Y^i - \Gamma^i_{jk}\dot{q}^j\dot{q}^k)\frac{\partial}{\partial v^i} + \dot{X}_1^i\frac{\partial}{\partial u^i} + \Bigg(\dot{X}_2^i - \frac{1}{2}\frac{\partial\Gamma^i_{jk}}{\partial q^\ell}\dot{q}^k\dot{q}^\ell X_1^j - \frac{1}{2}\frac{\partial\Gamma^i_{kj}}{\partial q^\ell}\dot{q}^k\dot{q}^\ell X_1^j - \frac{1}{2}(\Gamma^i_{jk} + \Gamma^i_{kj})(Y^k - \Gamma^k_{\ell m}\dot{q}^\ell\dot{q}^m)X_1^j - \frac{1}{2}(\Gamma^i_{jk} + \Gamma^i_{kj})\dot{q}^k\dot{X}_1^j\Bigg)\frac{\partial}{\partial w^i}.$$

A straightforward computation shows that this tangent vector field has the representation

$$c'(t) \oplus Y(t, c(t)) \oplus \big(\nabla_{c'(t)}X_1(t) + \tfrac{1}{2}T(X_1(t), c'(t))\big) \oplus \big(\nabla_{c'(t)}X_2(t) + \tfrac{1}{2}T(X_2(t), c'(t))\big)$$

which proves the lemma. ■

We may now prove our main result which relates the integral curves of $Z^T$ with solutions to the Jacobi equation.

**5.6 Theorem:** *Let $\nabla$ be an affine connection on $Q$ with $Z$ the corresponding geodesic spray. Let $c\colon I \to Q$ 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 $U, W\colon I \to TQ$ 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 $U$ and $W$ have the following properties:*

*(i) $U$ satisfies the Jacobi equation;*

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

Proof: Throughout the proof we represent points in $TTQ$ as the direct sum of tangent vectors to $Q$ using the connection on $\pi_{TQ}\colon TQ \to Q$ induced by $\nabla$. The tangent vector to the curve $t \mapsto U(t) \oplus W(t)$ at $t$ must equal $Z^T(U(t) \oplus W(t))$. By Lemma 5.5 and Proposition 5.4 this means that

$$\begin{aligned} \nabla_{c'(t)}U(t) &= W(t) - \tfrac{1}{2}T(U(t), c'(t)) \\ \nabla_{c'(t)}W(t) &= -\Omega(\mathrm{hlft}_{c'(t)}(U(t)), \mathrm{hlft}_{c'(t)}(c'(t))) - \tfrac{1}{2}T(W(t), c'(t)). \end{aligned} \tag{5.9}$$

The first of these equations proves (ii). To prove (i) we covariantly differentiate the first of equations (5.9). This yields, using the second of equations (5.9),

$$\begin{aligned} \nabla^2_{c'(t)}U(t) = -\Omega(\mathrm{hlft}_{c'(t)}(U(t)), \mathrm{hlft}_{c'(t)}(c'(t)))- \\ \tfrac{1}{2}T(W(t), c'(t)) - \tfrac{1}{2}\nabla_{c'(t)}(T(U(t), c'(t))). \end{aligned} \tag{5.10}$$

Now we see from Proposition 5.1 that

$$\begin{aligned} -\Omega(\mathrm{hlft}_{c'(t)}(U(t)), \mathrm{hlft}_{c'(t)}(c'(t))) = -R(U(t), c'(t))c'(t)- \\ \tfrac{1}{2}(\nabla_{c'(t)}T)(U(t), c'(t)) + \tfrac{1}{4}T(T(U(t), c'(t)), c'(t)). \end{aligned} \tag{5.11}$$

Combining (5.10), (5.11), and the first of equations (5.9) gives

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

which is simply the Jacobi equation, and so this proves (i). ■

**5.7 Remark:** Let us follow up on Remark 5.3–3 by showing that geodesics of $\nabla^T$ are indeed Jacobi fields. By Lemma 5.2 integral curves of $Z^T$ and of the geodesic spray for $\nabla^T$ are mapped to one another by the involution $I_Q$. Given

1. the representation $t \mapsto U(t) \oplus W(t)$ of integral curves of $Z^T$ as in Theorem 5.6,
2. the Ehresmann connection on $\pi_{TQ}\colon TQ \to Q$ described in Section 5.5, and
3. the coordinate expression for $I_Q$,

one verifies that the geodesics of $\nabla^T$ are exactly the vector fields $t \mapsto U(t)$ along geodesics as described in Theorem 5.6. But these are simply Jacobi fields according to the theorem. •

Now let's look at similar results relating to $Z^{T^*}$. First we give the coordinate formula for this vector field:

$$Z^{T^*} = v^i\frac{\partial}{\partial q^i} - \Gamma^i_{jk}v^jv^k\frac{\partial}{\partial v^i} + \frac{\partial\Gamma^\ell_{jk}}{\partial q^i}v^jv^k\beta_\ell\frac{\partial}{\partial\alpha_i} - \Big(\alpha_i - \Gamma^\ell_{ij}v^j\beta_\ell - \Gamma^\ell_{ji}v^j\beta_\ell\Big)\frac{\partial}{\partial\beta_i}. \tag{5.12}$$

To provide the decomposition for $Z^{T^*}$ we need an extra bit of notation. Fix $v_q \in T_qQ$ and note that for $u_q \in T_qQ$ we have $\Omega(\mathrm{hlft}_{v_q}(u_q), \mathrm{hlft}_{v_q}(v_q)) \in V_{v_q}TQ \simeq T_qQ$. Thus we may regard $u_q \mapsto \Omega(\mathrm{hlft}_{v_q}(u_q), \mathrm{hlft}_{v_q}(v_q))$ as an endomorphism of $T_qQ$. Let us denote the

dual endomorphism by $\beta_q \mapsto \Omega^*(\mathrm{vlft}_{v_q}(\beta_q), \mathrm{hlft}_{v_q}(v_q))$. The reason for this odd choice of notation for a dual endomorphism will become clear shortly.

We need more notation concerning the curvature and torsion tensors $R$ and $T$. For $u_q, v_q \in T_qQ$ and $\alpha_q \in T_q^*Q$, define $R^*(\alpha_q, u_q)v_q \in T_q^*Q$ by

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

and similarly define $T^*(\alpha_q, u_q) \in T_q^*Q$ by

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

With these tensors defined, we say that a one-form field $\alpha\colon I \to T^*Q$ along a geodesic $c\colon I \to Q$ of $\nabla$ is a solution of the ***adjoint Jacobi equation*** if

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

for $t \in I$.

Now let us recall the splittings associated with the connection $H(T^*TQ)$ on $\pi_{T^*TQ}\colon T^*TQ \to TQ$ which is described in Section 5.8. For $\Lambda_{v_q} \in T^*_{v_q}TQ$ we have

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

We then write $\Lambda_{v_q} \in T^*_{v_q}TQ$ as $\alpha_{v_q} \oplus \beta_{v_q}$ for some $\alpha_{v_q}, \beta_{v_q} \in T_q^*Q$, where we again make an abuse of notation. This then gives the following formula for $Z^{T^*}$ with respect to our splitting.

**5.8 Proposition:** $Z^{T^*}(\alpha_{v_q} \oplus \beta_{v_q}) = v_q \oplus 0 \oplus \big(\Omega^*(\mathrm{vlft}_{v_q}(\beta_{v_q}), \mathrm{hlft}_{v_q}(v_q)), -\alpha_{v_q}\big)$.

To demonstrate the relationship between integral curves of $Z^{T^*}$ and solutions to the adjoint Jacobi equation, we have the following analogue to Lemma 5.5.

**5.9 Lemma:** *Let $Y$ be a* $\mathrm{LIC}^\infty$ *vector field on $Q$ and suppose that $c\colon I \to Q$ is the* LAD *curve satisfying $\nabla_{c'(t)}c'(t) = Y(t, c(t))$, and denote by $\sigma\colon I \to TQ$ the tangent vector field of $c$ (i.e., $\sigma = c'$). Let $\Lambda\colon I \to T^*TQ$ be a* LAC *one-form field along $\sigma$, and denote $\Lambda(t) = \Lambda^1(t) \oplus \Lambda^2(t) \in T^*_{c(t)}Q \oplus T^*_{c(t)}Q \simeq T^*_{\sigma(t)}TQ$. Then the tangent vector to the curve $t \mapsto \Lambda(t)$ is given by $c'(t) \oplus Y(t, c(t)) \oplus \tilde{\Lambda}^1(t) \oplus \tilde{\Lambda}^2(t)$ where*

$$\tilde{\Lambda}^1(t) = \nabla_{c'(t)}\Lambda^1(t) - \tfrac{1}{2}T^*(\Lambda^1(t), c'(t)), \qquad \tilde{\Lambda}^2(t) = \nabla_{c'(t)}\Lambda^2(t) - \tfrac{1}{2}T^*(\Lambda^2(t), c'(t)).$$

Proof: In coordinates the curve $t \mapsto \Lambda(t)$ has the form

$$(q^i(t), \dot{q}^j(t), \Lambda^1_k(t) + \tfrac{1}{2}(\Gamma^m_{kr} + \Gamma^m_{rk})\dot{q}^r(t)\Lambda^2_m(t), \Lambda^2_i(t)).$$

The tangent vector to this curve is then given by

$$\dot{q}^i\frac{\partial}{\partial q^i} + (Y^i - \Gamma^i_{jk}\dot{q}^j\dot{q}^k)\frac{\partial}{\partial v^i} + \Bigg(\dot{\Lambda}^1_i + \frac{1}{2}\frac{\partial\Gamma^j_{ik}}{\partial q^\ell}\dot{q}^\ell\dot{q}^k\Lambda^2_j + \frac{1}{2}\frac{\partial\Gamma^j_{ki}}{\partial q^\ell}\dot{q}^\ell\dot{q}^k\Lambda^2_j + \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})(Y^k - \Gamma^k_{\ell m}\dot{q}^\ell\dot{q}^m)\Lambda^2_j + \frac{1}{2}(\Gamma^j_{ik} + \Gamma^j_{ki})\dot{q}^k\dot{\Lambda}^2_j\Bigg)\frac{\partial}{\partial\alpha_i} + \dot{\Lambda}^2_i\frac{\partial}{\partial\beta^i}.$$

A straightforward computation shows that this tangent vector field has the representation

$$c'(t) \oplus Y(t, c(t)) \oplus \big(\nabla_{c'(t)}\Lambda^1(t) - \tfrac{1}{2}T^*(\Lambda^1(t), c'(t))\big) \oplus \big(\nabla_{c'(t)}\Lambda^2(t) - \tfrac{1}{2}T^*(\Lambda^2(t), c'(t))\big)$$

which proves the lemma. ■

We may now prove our main result which relates the integral curves of $Z^T$ with solutions to the adjoint Jacobi equation.

**5.10 Theorem:** *Let $\nabla$ be an affine connection on $Q$ with $Z$ the corresponding geodesic spray. Let $c\colon I \to Q$ be a geodesic with $t \mapsto \sigma(t) \triangleq c'(t)$ the corresponding integral curve of $Z$. Let $a \in I$ and let $\theta, \lambda \in T^*_{c(a)}Q$, and define one-form fields $\Theta, \Lambda\colon I \to T^*Q$ along $c$ by asking that $t \mapsto \Theta(t) \oplus \Lambda(t) \in T^*_{c(t)}Q \oplus T^*_{c(t)}Q \simeq T^*_{\sigma(t)}TQ$ be the integral curve of $Z^{T^*}$ with initial conditions $\theta \oplus \lambda \in T^*_{c(a)}Q \oplus T^*_{c(a)}Q \simeq T^*_{\sigma(a)}TQ$. Then $\Theta$ and $\Lambda$ have the following properties:*

*(i) $\Lambda$ satisfies the adjoint Jacobi equation;*

*(ii) $\Theta(t) = -\nabla_{c'(t)}\Lambda(t) + \frac{1}{2}T^*(\Lambda(t), c'(t))$.*

Proof: Throughout the proof we represent points in $T^*TQ$ as the direct sum of cotangent vectors to $Q$ using the connection on $\pi_{T^*Q}\colon T^*Q \to Q$ induced by $\nabla$. The tangent vector to the curve $t \mapsto \Theta(t) \oplus \Lambda(t)$ at $t$ must equal $Z^{T^*}(\Theta(t) \oplus \Lambda(t))$. By Lemma 5.9 and Proposition 5.8 this means that

$$\begin{aligned} \nabla_{c'(t)}\Theta(t) &= \Omega^*(\mathrm{vlft}_{c'(t)}(\Lambda(t)), \mathrm{hlft}_{c'(t)}(c'(t))) + \tfrac{1}{2}T^*(\Theta(t), c'(t)) \\ \nabla_{c'(t)}\Lambda(t) &= -\Theta(t) + \tfrac{1}{2}T^*(\Lambda(t), c'(t)). \end{aligned} \tag{5.13}$$

The second of these equations proves (i). To prove (ii) we covariantly differentiate the second of equations (5.13). This yields, using the first of equations (5.13),

$$\begin{aligned} \nabla^2_{c'(t)}\Lambda(t) = -\Omega^*(\mathrm{hlft}_{c'(t)}(\Lambda(t)), \mathrm{hlft}_{c'(t)}(c'(t))) - \tfrac{1}{2}T^*(\Theta(t), c'(t)) + \\ \tfrac{1}{2}(\nabla_{c'(t)}T^*)(\Lambda(t), c'(t))) + \tfrac{1}{2}T^*(\nabla_{c'(t)}\Lambda(t), c'(t)). \end{aligned} \tag{5.14}$$

Now we see from Proposition 5.1 that

$$\begin{aligned} \Omega^*(\mathrm{vlft}_{c'(t)}(\Lambda(t)), \mathrm{hlft}_{c'(t)}(c'(t))) = R^*(\Lambda(t), c'(t))c'(t) + \\ \tfrac{1}{2}(\nabla_{c'(t)}T^*)(\Lambda(t), c'(t)) - \tfrac{1}{4}T^*(T^*(\Lambda(t), c'(t)), c'(t)). \end{aligned} \tag{5.15}$$

Combining (5.14), (5.15), and the second of equations (5.13) gives

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

which is simply the adjoint Jacobi equation, and so this proves (i). ■

**5.11 Remarks:** 1. Note that the adjoint Jacobi equation, along with the geodesic equations themselves, of course, contain the non-trivial dynamics of the Hamiltonian vector field $Z^{T^*}$. Note also that the Hamiltonian in our splitting of $T^*TQ$ is simply given by $\alpha_{v_q} \oplus \beta_{v_q} \mapsto \alpha_{v_q} \cdot v_q$. Thus, while the Hamiltonian assumes a simple form in this splitting, evidently the symplectic form becomes rather complicated. However, since the maximum principle employs the Hamiltonian in its statement, the simple form of the Hamiltonian will be very useful for us.

2. We may express the content of Proposition 3.4, in the case when the vector field in question is the geodesic spray, as follows. We use the notation of Propositions 5.4 and 5.8. Let $u_{v_q} \oplus w_{v_q} \in T_{v_q}TQ$ and $\alpha_{v_q} \oplus \beta_{v_q} \in T^*_{v_q}TQ$. Let $\mathrm{ver}(Z^T(u_{v_q} \oplus w_{v_q}))$ be the vertical part of $Z^T(u_{v_q} \oplus w_{v_q})$ which we think of as a vector in $T_qQ \oplus T_qQ$. In a similar manner we think of $\mathrm{ver}(Z^{T^*}(\alpha_{v_q} \oplus \beta_{v_q}))$ as a vector in $T^*_qQ \oplus T^*_qQ$. A straightforward computation shows that

$$\big\langle \mathrm{ver}(Z^{T^*}(\alpha_{v_q} \oplus \beta_{v_q})); u_{v_q} \oplus w_{v_q} \big\rangle + \big\langle \alpha_{v_q} \oplus \beta_{v_q}; \mathrm{ver}(Z^T(u_{v_q} \oplus w_{v_q})) \big\rangle = 0. \qquad \bullet$$

The Jacobi equation and the adjoint Jacobi equation have a closer relationship when $\nabla$ is the Levi-Civita connection $\overset{g}{\nabla}$ associated to a Riemannian metric $g$. We denote by $g^\flat\colon TQ \to T^*Q$ the vector bundle map associated with the symmetric $(0,2)$ tensor $g$. Since $g$ is nondegenerate, there is an induced vector bundle metric on $T^*Q$ which we denote by $g^{-1}$. We denote by $g^\sharp\colon T^*Q \to TQ$ the vector bundle map associated with the symmetric $(2,0)$ tensor $g^{-1}$. Since $g$ is nondegenerate, $g^\flat$ and $g^\sharp$ are inverses of one another. We have the following result.

**5.12 Proposition:** *Let $g$ be a Riemannian metric on $Q$ with $\overset{g}{\nabla}$ the Levi-Civita affine connection. If $c\colon I \to Q$ is a geodesic of $\overset{g}{\nabla}$ then a one-form field $\lambda\colon I \to T^*Q$ along $c$ is a solution of the adjoint Jacobi equation if and only if the vector field $g^\sharp \circ \lambda$ along $c$ is a solution of the Jacobi equation.*

Proof: Using the fact that $\overset{g}{\nabla} g = 0$ we compute

$$\nabla^2_{c'(t)} g^\sharp(\lambda(t)) = g^\sharp(\nabla^2_{c'(t)}\lambda(t)). \tag{5.16}$$

Now, using (4.4), and for $u \in T_{c(t)}Q$ we compute

$$\begin{aligned}\langle R^*(\lambda(t), c'(t))c'(t); u\rangle &= \langle \lambda(t); R(u, c'(t))c'(t)\rangle \\ &= g(R(u, c'(t))c'(t), g^\sharp(\lambda(t))) \\ &= g(R(g^\sharp(\lambda(t)), c'(t))c'(t), u) \\ &= \langle g^\flat(R(g^\sharp(\lambda(t)), c'(t))c'(t)); u\rangle.\end{aligned}$$

This implies that

$$g^\sharp(R^*(\lambda(t), c'(t))c'(t)) = R(g^\sharp(\lambda(t)), c'(t))c'(t). \tag{5.17}$$

Combining (5.16) and (5.17) gives

$$\nabla^2_{c'(t)} g^\sharp(\lambda(t)) + R(g^\sharp(\lambda(t)), c'(t))c'(t) = g^\sharp\big(\nabla^2_{c'(t)}\lambda(t) + R^*(\lambda(t), c'(t))c'(t)\big)$$

and the result now follows since $\overset{g}{\nabla}$ is torsion-free and since $g^\sharp$ is a vector bundle isomorphism. ■

**5.10. Representation of $(\mathrm{vlft}(X))^{T^*}$.** We complete this section by providing the formula for $(\mathrm{vlft}(X))^{T^*}$ for a vector field $X$ on $Q$. To give this formula, we need some notation. For $\alpha_q \in T_qQ$, define $(\nabla X)^*(\beta_q) \in T_q^*Q$ by

$$\langle (\nabla X)^*(\alpha_q); u_q\rangle = \langle \alpha_q; \nabla_{u_q} X\rangle.$$

We now adopt the same notation as used in Proposition 5.8. The proof is accomplished easily in coordinates using the coordinate formula (3.3).

**5.13 Proposition:** *If $X$ is a vector field on $Q$ then*

$$(\mathrm{vlft}(X))^{T^*}(\alpha_{v_q} \oplus \beta_{v_q}) = 0 \oplus X(q) \oplus \big(\tfrac{1}{2}T^*(\beta_{v_q}, X(q)) - (\nabla X)^*(\beta_{v_q})\big) \oplus 0.$$

**5.14 Remark:** This representation of $(\mathrm{vlft}(X))^{T^*}$ has a further geometric interpretation as follows. Let $X^{T^*}$ be the cotangent lift of $X$ to a vector field on $T^*Q$. This vector field may then be written with respect to the decomposition corresponding to the connection on $\pi_{T^*Q}\colon T^*Q \to Q$ given in Section 5.6. If we do so, we have

$$X^{T^*}(\alpha_q) = X(q) \oplus \big(\tfrac{1}{2}T^*(\alpha_q, X(q)) - (\nabla X)^*(\alpha_q)\big).$$

The interested reader will see that this is consistent with our explanation in Section 3.4 of the relationship between $(\mathrm{vlft}(X))^{T^*}$ and $\mathrm{vlft}(X^{T^*})$. •

# Part II. Optimal Control

## 6. The maximum principle for control affine systems on manifolds

In our geometric setting, it is essential to be able to talk about the maximum principle on manifolds, rather than on open subsets of Euclidean space. To do this, we employ the framework of Sussmann [1998], although we simplify the framework there to a large degree by assuming smoothness of the problem data (along with other simplifications).

**6.1. Control affine systems.** Let us first deal with generalities concerning control affine systems. A ***control affine system*** is a triple $\Sigma = (M, \mathscr{F}, U)$ where $M$ is an $n$-dimensional, separable, connected, Hausdorff manifold, $\mathscr{F} = \{f_0, f_1, \dots, f_m\}$ is a collection of $C^\infty$ vector fields on $M$, and $U\colon M \to \mathbf{2}^{\mathbb{R}^m}$ assigns a subset of $\mathbb{R}^m$ to each point $x \in M$. We denote by $U_x \subset \mathbb{R}^m$ the image of $x \in M$ under $U$. We call $U$ the ***input range map***. We say that $U$ is ***constant*** if there exists a subset $S$ of $\mathbb{R}^m$ so that $U_x = S$ for every $x \in M$. Roughly speaking we deal with the control system

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

on $M$ where the control function $u$ satisfies $u(t) \in U_{x(t)}$. It will pay to be a bit more precise about the nature of the control problem if we are to make a rigorous statement of the maximum principle.

**6.1 Remark:** We remark that our use of control sets which depend on state is nonstandard, but adds nothing to the complication of the control problems we consider. Although it seems like a strange generalisation, it is sensible since it is not reasonable to expect the possible values for the controls to be generally independent of state. It is also useful to note that our systems have the "compact substitution property" in the terminology of Sussmann [1998]. This means that his general results are applicable to our systems. •

We shall now specify the exact nature of the controls for such a system. Recall [see, e.g., Sontag 1998] that if $u\colon I \to U$ is a measurable map into a separable metric space $U$, and if $f\colon U \times M \to TM$ is a continuous map which is $C^\infty$ in the second argument and is such that $f(u_0, x) \in T_xM$ for each $u_0 \in U$, then the map $(t, x) \mapsto f(u(t), x)$ is a LIC$^\infty$ Carathéodory vector field. Thus, given a measurable map $u\colon I \to \mathbb{R}^m$ we may define a LIC$^\infty$ Carathéodory vector field on $I \times M$ by

$$\mathscr{F}_u(t, x) = f_0(x) + u^a(t) f_a(x).$$

A ***controlled trajectory*** for a control affine system $\Sigma = (M, \mathscr{F}, U)$ is a pair $\gamma = (u, c)$ where $u\colon I \to \mathbb{R}^m$ is measurable, $c\colon I \to M$ is an integral curve for the LIC$^\infty$ vector field $\mathscr{F}_u$, and where $u(t) \in U_{c(t)}$ for almost every $t \in I$. If in this definition the interval $I$ is compact, then we call $\gamma$ a ***controlled arc***. We denote by $\mathrm{Ctraj}(\Sigma)$ the set of controlled trajectories for $\Sigma$, and by $\mathrm{Carc}(\Sigma)$ the set of controlled arcs.

**6.2. An optimal control problem.** We now consider a control affine system $\Sigma = (M, \mathscr{F}, U)$ as above, but now we make the addition of a ***cost function***[3] $F$ for $\Sigma$ which is defined on

$$\mathscr{D}(\Sigma) = \{(u, x) \in \mathbb{R}^m \times M \mid u \in U_x\}.$$

Although we are only interested in the value of $F$ on this set, it is convenient to suppose $F$ to be the restriction of a continuous function from $\mathbb{R}^m \times M$ to $\mathbb{R}$ and for which $x \mapsto F(u, x)$ is of class $C^\infty$ for every $u \in \mathbb{R}^m$. Note that the function $(t, x) \mapsto F(u(t), x)$ is a LIC$^\infty$ Carathéodory function if $\gamma = (u, c) \in \mathrm{Ctraj}(\Sigma)$. We shall say that $\gamma = (u, c) \in \mathrm{Ctraj}(\Sigma)$ is ***F-acceptable*** if the function $t \mapsto F(u(t), c(t))$ is locally integrable. We denote by $\mathrm{Ctraj}(\Sigma, F)$ the subset of $\mathrm{Ctraj}(\Sigma)$ consisting of $F$-acceptable controlled trajectories, and by $\mathrm{Carc}(\Sigma, F)$ the set of $F$-acceptable controlled arcs.

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

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

Let $S_0$ and $S_1$ be disjoint submanifolds of $M$. We denote by

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

In like fashion, for $a, b \in \mathbb{R}$ with $a < b$, we define

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

The problems concerning the optimal path connecting two submanifolds are stated as follows.

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

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

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

A special case of this problem occurs when $S_0 = \{x_0\}$ and $S_1 = \{x_1\}$ for two points $x_0, x_1 \in M$.

[3] Sussmann [1998] calls such functions ***Lagrangians***. However, since we are dealing with control systems whose dynamics themselves are sometimes Lagrangian, we refrain from using this notation as it might lead to one Lagrangian too many.

**6.3. The maximum principle.** Now that we have stated clearly the optimal control problem we wish to investigate, let us state necessary conditions for solutions of this problem. Key to the necessary conditions of the maximum principle is the use of a Hamiltonian formalism. Let $\Sigma = (M, \mathscr{F}, U)$ be a control affine system and let $F$ be a cost function for $\Sigma$. We define a subset $\mathscr{D}^*(\Sigma)$ of $U \times T^*M$ by

$$\mathscr{D}^*(\Sigma) = \{(u, \alpha_x) \in U \times T^*M \mid u \in U_x\}.$$

We define the ***Hamiltonian*** $H^{\Sigma,F}$ as a function on $\mathscr{D}^*(\Sigma)$ by

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

From this we define the ***minimum Hamiltonian*** as the function on $T^*M$ defined by

$$H^{\Sigma,F}_{\min}(\alpha_x) = \inf_{u \in U_x} H^{\Sigma,F}(u, \alpha_x)$$

and we adopt the notation $H^{\Sigma,F}_{\min}(\alpha_x) = -\infty$ if for each $C \in \mathbb{R}$ there exists $u \in U_x$ so that $H^{\Sigma,F}(u, \alpha_x) < C$. If $u \in U_x$ has the property that $H^{\Sigma,F}(u, \alpha_x) = H^{\Sigma,F}_{\min}(\alpha_x)$, then we say $H^{\Sigma,F}_{\min}$ is ***realised*** by $u$ at $\alpha_x$. If $\gamma = (u, c) \in \mathrm{Ctraj}(\Sigma, F)$ then the function $H^{\Sigma,F}_u(t, \alpha_x) \mapsto H^{\Sigma,F}(u(t), \alpha_x)$ is $\mathrm{LIC}^\infty$. Therefore, corresponding to this function will be a $\mathrm{LIC}^\infty$ Hamiltonian vector field $X_{H^{\Sigma,F}_u}$ on $T^*M$.

If $\gamma = (u, c) \in \mathrm{Ctraj}(\Sigma, F)$ with $u$ and $c$ defined on the interval $I$, then a LAC one-form field $\chi\colon I \to T^*M$ along $c$ is ***minimising for*** $\boldsymbol{(\Sigma, F)}$ ***along*** $\boldsymbol{u}$ if

$$H^{\Sigma,F}(u(t), \chi(t)) \le H^{\Sigma,F}_{\min}(\chi(t))$$

for almost every $t \in I$.

We now state the maximum principle as we shall employ it.

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

*(i) $\chi(a) \in \mathrm{ann}(T_{c(a)}S_0)$ and $\chi(b) \in \mathrm{ann}(T_{c(b)}S_1)$;*

*(ii) $t \mapsto \chi(t)$ is an integral curve of $X_{H^{\Sigma,\chi_0 F}_u}$;*

*(iii) $\chi$ is minimising for $(\Sigma, \chi_0 F)$ along $u$;*

*(iv) either $\chi_0 = 1$ or $\chi(a) \neq 0$.*

*If $U$ is constant then we have*

*(v) there exists a constant $C \in \mathbb{R}$ so that $H^{\Sigma,F}(u(t), \chi(t)) = C$ a.e.*

*If $U$ is constant and if $\gamma = (u, c)$ is a solution of $\mathscr{P}(\Sigma, F, S_0, S_1)$, then condition (v) can be replaced with*

*(vi) $H^{\Sigma,F}(u(t), \chi(t)) = 0$ a.e.*

**6.4 Remarks:** 1. The maximum principle we state is a specific case of that stated by Sussmann [1998]. In his paper, Sussmann considers a more general class of control systems where the control vector fields, as well as the cost function, may only be locally integrally Lipschitz. It is possible that for certain applications one may wish to allow nonsmooth data, but we will not be considering such problems in this work.

2. If $S_0 = \{x_0\}$ then $\chi(a)$ is unrestricted (modulo requirement (iv)). Similarly, if $S_1 = \{x_1\}$ then $\chi(b)$ is unrestricted.
3. Since the Hamiltonian vector field $X_{H_u^{\Sigma,F}}$ is linear in the fibre variables (Remark 3.2–2), the condition (iv) in the statement of the maximum principle asserts that $(\chi_0, \chi(t))$ will be non-zero for $t \in [a, b]$.
4. Note that for either of conditions (v) or (vi) to hold, we cannot allow the input range map to be arbitrary. In the language of Sussmann [1998], our condition that $U$ be constant implies that $(\Sigma, L)$ is "closed under interval insertions and deletions." •

## 7. The maximum principle on manifolds with an affine connection

We now reap the benefits of the work in the first part of the paper to provide a concise translation of the maximum principle for systems whose drift vector field is the geodesic spray associated with an affine connection, and whose control vector fields are vertically lifted vector fields.

**7.1. Affine connection control systems.** To be precise about the control systems we are considering, we define an ***affine connection control system*** to be a quadruple $\Sigma_{\text{aff}} = (Q, \nabla, \mathscr{Y}, U)$ where $Q$ is a smooth, finite-dimensional, separable, Hausdorff manifold (the configuration space), $\nabla$ is a smooth affine connection on $Q$, $\mathscr{Y} = \{Y_1, \dots, Y_m\}$ is collection of smooth vector fields on $Q$, and $U\colon Q \to \mathbf{2}^{\mathbb{R}^m}$ is a map into the set of subsets of $\mathbb{R}^m$. As with control affine systems, we denote $U_q = U(q) \subset \mathbb{R}^m$, and we also say that $U$ is ***constant*** if there exists a subset $S$ of $\mathbb{R}^m$ with the property that $U_q = S$ for all $q \in Q$. The control systems we consider have the form

$$\nabla_{c'(t)} c'(t) = u^a(t) Y_a(c(t)). \tag{7.1}$$

In order to make this a control affine system, the state space should be $TQ$ since (7.1) is a control system of second-order. One verifies that the proper first-order equation on $TQ$ is

$$\dot{v}(t) = Z(v(t)) + u^a(t)\,\mathrm{vlft}(Y_a)(v(t)).$$

where $Z$ is the geodesic spray of $\nabla$. Thus we obtain from an affine connection control system $\Sigma_{\text{aff}} = (Q, Z, \mathscr{Y}, U)$ a control affine system[4] $\Sigma = (TQ, \{Z, \mathrm{vlft}(Y_1), \dots, \mathrm{vlft}(Y_m)\}, U^T)$ where $U^T\colon TQ \to \mathbf{2}^{\mathbb{R}^m}$ is defined by $U^T(v_q) = U(q)$.

**7.1 Remark:** We observe that all the problem data for an affine connection control system is defined on the configuration manifold $Q$. When we render this system a control affine system, we do so by lifting the problem data to $TQ$ in a natural manner. In doing so, we allow the application of all the machinery which has been developed for control affine

[4]The dual uses of the word "affine" appears to be unavoidable here, unfortunately.

systems. However, one is within bounds of reason to expect the answers obtained to be expressible in terms of the original data on $Q$. For a certain type of controllability, this program is carried out by Lewis and Murray [1997]. Here we carry this out for the formulation of the maximum principle. •

Note that if $u\colon I \to \mathbb{R}^m$ is measurable then the map $Z_u^{\mathscr{Y}} : I \times TQ \to TTQ$ given by

$$Z_u^{\mathscr{Y}}(t, v_q) = Z(v_q) + u^a(t)\,\mathrm{vlft}(Y_a(q))$$

defines a $\mathrm{LIC}^\infty$ second-order Carathéodory vector field. Therefore, this vector field will possess an integral curve $\sigma\colon J \to TQ$, for $J \subset I$ sufficiently small, through each $v_q \in TQ$. Therefore there will exist a LAD curve $c\colon J \to Q$ with the property that $c'(t) = \sigma(t)$ for every $t \in J$ and which passes through $q$ with initial velocity $v_q$. A ***controlled trajectory*** for $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ is a pair $\gamma = (u, c)$ where $u\colon I \to \mathbb{R}^m$ is measurable, $c\colon I \to Q$ is LAD with $c'$ being an integral curve of $Z_u^{\mathscr{Y}}$, and $u(t) \in U_{c(t)}$ for almost every $t \in I$. If the interval $I$ is compact then $\gamma$ is called a ***controlled arc***. We denote by $\mathrm{Ctraj}(\Sigma_{\mathrm{aff}})$ the set of controlled trajectories of $\Sigma_{\mathrm{aff}}$, and by $\mathrm{Carc}(\Sigma_{\mathrm{aff}})$ the set of controlled arcs.

Let us state a property of affine connection control systems which can help to determine the reasonableness of an optimal control problem.

**7.2 Lemma:** *Let $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be an affine connection control system with $(u, c) \in \mathrm{Ctraj}(\Sigma_{\mathrm{aff}})$ defined on $I \subset \mathbb{R}$. For $\lambda > 0$ define $I_\lambda = \{\frac{1}{\lambda}t \mid t \in I\}$, and define $\tilde{u}\colon I_\lambda \to \mathbb{R}^m$ and $\tilde{c}\colon I_\lambda \to Q$ by $\tilde{u}(t) = \lambda^2 u(\lambda t)$ and $\tilde{c}(t) = c(\lambda t)$. Then, if $\tilde{u}(t) \in U_{\tilde{c}(t)}$ for $t \in I_\lambda$, then $(\tilde{u}, \tilde{c}) \in \mathrm{Ctraj}(\Sigma_{\mathrm{aff}})$.*

Proof: We compute

$$\nabla_{\tilde{c}'(t)}\tilde{c}'(t) = \lambda^2 \nabla_{c'(\lambda t)} c'(\lambda t) = \lambda^2 u(\lambda t) = \tilde{u}(t). \qquad \blacksquare$$

**7.2. Optimal control problems for affine connection control systems.** Of course, since an affine connection control system defines a control affine system, one may simply formulate an optimal control problem on $TQ$ exactly as was done in Section 6.2. However, we wish to choose a class of cost functions which reflects the fact that the problem data for an affine connection control system is defined on $Q$.

Let $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be an affine connection control system. We first define

$$\mathscr{D}_Q(\Sigma_{\mathrm{aff}}) = \{(u, q) \in \mathbb{R}^m \times Q \mid u \in U_q\}.$$

A ***$\mathbb{R}^m$-dependent $(0, r)$ tensor field*** on $Q$ is a map $A\colon \mathbb{R}^m \times Q \to \mathrm{T}_r^0(TQ)$ such that

DTF1. $A$ is continuous, and

DTF2. $q \mapsto A(u, q)$ is a smooth $(0, r)$ tensor field for every $u \in \mathbb{R}^m$.

Note that we are really only interested in the value of $\mathbb{R}^m$-dependent tensor fields when evaluated at points $(u, q) \in \mathscr{D}_Q(\Sigma_{\mathrm{aff}})$. However, for simplicity we suppose them to be defined on all of $\mathbb{R}^m$. We let $r \geq 0$ and let $A\colon \mathscr{D}_Q(\Sigma_{\mathrm{aff}}) \to \mathrm{T}_r^0(TQ)$ be a $\mathbb{R}^m$-dependent

symmetric $(0,r)$ tensor field on $Q$.[5] We let $f\colon \mathbb{R}\to\mathbb{R}$ be a class $C^\infty$ function[6] and let $\mathscr{A}=(A,f)$ and define

$$\mathscr{D}_{TQ}(\Sigma_{\text{aff}})=\{(u,v_q)\in\mathbb{R}^m\times TQ \mid\ u\in U_q\}.$$

A ***cost function*** for $\Sigma_{\text{aff}}$ is a function $F_{\mathscr{A}}\colon \mathscr{D}_{TQ}(\Sigma_{\text{aff}})\to\mathbb{R}$ of the form

$$F_{\mathscr{A}}(u,v_q)=f(A(u,q)(v_q,\dots,v_q)).$$

We shall need some notation involving symmetric $(0,r)$ tensor fields. For $v_1,\dots,v_{r-1}\in T_qQ$ we define $\hat{A}(v_1,\dots,v_{r-1})\in T_q^*Q$ by

$$\langle\hat{A}(v_1,\dots,v_{r-1});u\rangle=A(u,v_1,\dots,v_{r-1}),\quad u\in T_qQ.$$

We adopt the convention that if $A$ is a $(0,0)$ tensor field (i.e., $A$ is a function) then $\hat{A}=0$. Obviously this notation extends to tensor fields which are $\mathbb{R}^m$-dependent. The following lemma provides the form of a certain Hamiltonian vector field which we will encounter.

**7.3 Lemma:** *Let $A$ be a symmetric $(0,r)$ tensor field on $Q$, let $f\colon\mathbb{R}\to\mathbb{R}$ be of class $C^\infty$, and in the splitting of $T^*TQ$ defined in Section 5.8 consider a function defined by*

$$\alpha_{v_q}\oplus\beta_{v_q}\mapsto f(A(v_q,\dots,v_q)).$$

*The Hamiltonian vector field on $T^*TQ$ generated by this function has the decomposition*

$$f'(A(v_q,\dots,v_q))\big(0\oplus 0\oplus\big(-\nabla A(v_q,\dots,v_q)-\tfrac{r}{2}T^*(\hat{A}(v_q,\dots,v_q),v_q)\big)\oplus(-r\hat{A}(v_q,\dots,v_q))\big).$$

Proof: In natural coordinates for $T^*TQ$ the Hamiltonian function defined in the lemma is given by

$$((q,v),(\alpha,\beta))\mapsto f(A_{j_1\cdots j_r}v^{j_1}\cdots v^{j_r}).$$

The corresponding Hamiltonian vector field in natural coordinates is given by

$$f'(A_{j_1\cdots j_r}v^{j_1}\cdots v^{j_r})\Big(-\frac{\partial A_{j_1\cdots j_r}}{\partial q^i}v^{j_1}\cdots v^{j_r}\frac{\partial}{\partial\alpha_i}-rA_{ij_2\cdots j_r}v^{j_2}\cdots v^{j_r}\frac{\partial}{\partial\beta_i}\Big).$$

We now express this in a basis adapted to the splitting of $T_{\Lambda_{v_q}}T^*TQ$ to get

$$f'(A_{j_1\cdots j_r}v^{j_1}\cdots v^{j_r})\Bigg(-\Bigg(\frac{\partial A_{j_1\cdots j_r}}{\partial q^i}v^{j_1}\cdots v^{j_r}-\frac{r}{2}(\Gamma^\ell_{ik}+\Gamma^\ell_{ki})v^kA_{\ell j_2\cdots j_r}v^{j_2}\cdots v^{j_r}\Bigg)\frac{\partial}{\partial\alpha_i}-$$
$$rA_{ij_2\cdots j_r}v^{j_2}\cdots v^{j_r}\Big(\frac{\partial}{\partial\beta_i}+\frac{1}{2}(\Gamma^i_{k\ell}+\Gamma^i_{\ell k})v^\ell\frac{\partial}{\partial\alpha_k}\Big)\Bigg).$$

From this we see that the representation of the Hamiltonian vector field is

$$\alpha_{v_q}\oplus\beta_{v_q}\mapsto$$
$$f'(A(v_q,\dots,v_q))\big(0\oplus 0\oplus\big(-\nabla A(v_q,\dots,v_q)-\tfrac{r}{2}T^*(\hat{A}(v_q\dots,v_q),v_q)\big)\oplus(-r\hat{A}(v_q,\dots,v_q))\big).$$

This completes the proof. ■

[5] A $(0,r)$ tensor field $\tau$ is ***symmetric*** if $\tau(v_{\sigma(1)},\dots,v_{\sigma(r)})=\tau(v_1,\dots,v_r)$ for every $\sigma$ in the permutation group $S_r$ on $r$ symbols, for every $v_1,\dots,v_r\in T_qQ$, and for every $q\in Q$.

[6] The smoothness assumption may be relaxed here following Sussmann [1998].

The lemma clearly extends to $\mathbb{R}^m$-dependent tensor fields.

This is an appropriate setting in which to present the following lemma, although it will not be used until Section 8. As we talk about symmetric $(0,r)$ tensor fields, we may also talk about symmetric $(r,0)$ tensor fields. And we generate for these latter some notation similar to that generated for the former. To wit, if $B$ is a symmetric $(r,0)$ tensor field and if $\alpha^1,\dots,\alpha^{r-1}\in T_q^*Q$ we define $\hat{B}(\alpha^1,\dots,\alpha^{r-1})\in T_qQ$ by

$$\langle\beta;\hat{B}(\alpha^1,\dots,\alpha^{r-1})\rangle = B(\beta,\alpha^1,\dots,\alpha^{r-1}),\quad \beta\in T_q^*Q.$$

We now state the lemma.

**7.4 Lemma:** *Let $B$ be a symmetric $(r,0)$ tensor field on $Q$ and define a function on $T^*TQ$ by*

$$\alpha_{v_q}\oplus\beta_{v_q}\mapsto B(\beta_{v_q},\dots,\beta_{v_q}).$$

*The Hamiltonian vector field generated by this function has the representation*

$$0\oplus(r\hat{B}(\beta_{v_q},\dots,\beta_{v_q}))\oplus\big(\nabla B(\beta_{v_q},\dots,\beta_{v_q})-\tfrac{r}{2}T^*(\beta_{v_q},\hat{B}(\beta_{v_q},\dots,\beta_{v_q}))\big)\oplus 0.$$

Proof: In natural coordinates for $T^*TQ$ the function in the lemma has the form

$$((q,v),(\alpha,\beta))\mapsto B^{j_1\cdots j_r}\beta_{j_1}\cdots\beta_{j_r}.$$

Thus the Hamiltonian vector field associated with this function is given in natural coordinates by

$$rB^{ij_2\cdots j_r}\beta_{j_2}\cdots\beta_{j_r}\frac{\partial}{\partial v^i}-\frac{\partial B^{j_1\cdots j_r}}{\partial q^i}\beta_{j_1}\cdots\beta_{j_r}\frac{\partial}{\partial\alpha_i}.$$

If we write this in the splitting of $T_{\Lambda_{v_q}}T^*TQ$ we obtain the decomposition of the Hamiltonian vector field as

$$0\oplus(r\hat{B}(\beta_{v_q},\dots,\beta_{v_q}))\oplus\big(\nabla B(\beta_{v_q},\dots,\beta_{v_q})-\tfrac{r}{2}T^*(\beta_{v_q},\hat{B}(\beta_{v_q},\dots,\beta_{v_q}))\big)\oplus 0,$$

just as we have asserted. ■

Let $F_{\mathscr{A}}$ be a cost function for $\Sigma_{\text{aff}}$ as defined above. We say that $\gamma=(u,c)\in\text{Ctraj}(\Sigma)$ is ***$F_{\mathscr{A}}$-acceptable*** if the function $t\mapsto F_{\mathscr{A}}(u(t),c'(t))$ is locally integrable. We denote by $\text{Ctraj}(\Sigma_{\text{aff}},F_{\mathscr{A}})$ the set of $F_{\mathscr{A}}$-acceptable control trajectories, and by $\text{Carc}(\Sigma_{\text{aff}},F_{\mathscr{A}})$ the $F_{\mathscr{A}}$-acceptable control arcs.

For $\gamma=(u,c)\in\text{Carc}(\Sigma_{\text{aff}},F_{\mathscr{A}})$ where $u$ and $c$ are defined on $[a,b]$, we define

$$J^{\Sigma_{\text{aff}},F_{\mathscr{A}}}(\gamma)=\int_a^b F_{\mathscr{A}}(u(t),c'(t))\,\mathrm{d}t.$$

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 denote

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

For fixed $a, b \in \mathbb{R}$ with $a < b$ we define

$$\mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1}, [a,b]) = \{\gamma = (u,c) \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}) \mid \text{ where } u \text{ and } c \text{ are defined on } [a,b] \text{ and } c'(a) = v_{q_0} \text{ and } c'(b) = v_{q_1}\}.$$

The above subsets of controlled arcs correspond to fixing an initial and final configuration *and* velocity. For affine connection control systems, it also makes sense to consider only fixing the initial and final configuration while leaving the velocities free. Thus we define

$$\mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1) = \{\gamma = (u,c) \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}) \mid c(a) = q_0 \text{ and } c(b) = q_1 \text{ where } u \text{ and } c \text{ are defined on } [a,b] \text{ for some } a, b \in \mathbb{R}\},$$

and for fixed $a, b \in \mathbb{R}$ with $a < b$ we define

$$\mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1, [a,b]) = \{\gamma = (u,c) \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}) \mid \text{ where } u \text{ and } c \text{ are defined on } [a,b] \text{ and } c(a) = q_0 \text{ and } c(b) = q_1\}.$$

Now we define the control problems we consider.

**7.5 Definition:** Let $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be an affine connection control system, let $F_{\mathscr{A}}$ be a cost function for $\Sigma_{\mathrm{aff}}$, let $q_0, q_1 \in Q$, and let $v_{q_0} \in T_{q_0}Q$ and $v_{q_1} \in T_{q_1}Q$.

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

(ii) A controlled arc $\gamma_* \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1}, [a,b])$ is a ***solution of*** $\boldsymbol{\mathscr{P}_{[a,b]}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1})}$ if $J^{\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}}(\gamma_*) \le J^{\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}}(\gamma)$ for every $\gamma \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1}, [a,b])$.

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

(iv) A controlled arc $\gamma_* \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1, [a,b])$ is a ***solution of*** $\boldsymbol{\mathscr{P}_{[a,b]}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1)}$ if $J^{\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}}(\gamma_*) \le J^{\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}}(\gamma)$ for every $\gamma \in \mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1, [a,b])$. •

**7.6 Remark:** Lemma 7.2 immediately eliminates some possible optimal problems from consideration. For example, if the input range map $U$ is given by $U_q = \mathbb{R}^m$ for each $q \in Q$, then the time-optimal problem of steering between configurations does not make sense. This is because Lemma 7.2 implies that in this case, by using large controls, one may make the time of transit from $q_0$ to $q_1$ as small as one likes, providing one can find *some* controlled arc joining the points. •

**7.3. The maximum principle for affine connection control systems.** Before stating the maximum principle for the systems we are investigating, let us look at the Hamiltonian for these systems. In doing so, we will use the splitting of $T^*TQ$ which we presented in Proposition 5.8. Thus we write $\Lambda_{v_q} \in T^*_{v_q}TQ$ as $\alpha_{v_q} \oplus \beta_{v_q}$ for some appropriately defined $\alpha_{v_q}, \beta_{v_q} \in T^*_qQ$. We define

$$\mathscr{D}^*_{TQ}(\Sigma_{\mathrm{aff}}) = \{(u, \Lambda_{v_q}) \in \mathbb{R}^m \times T^*TQ \mid u \in U_q\}.$$

The ***Hamiltonian*** for an affine connection control system $\Sigma_{\text{aff}}$ with cost function $F_{\mathscr{A}}$ is the function on $\mathscr{D}^*_{TQ}(\Sigma_{\text{aff}})$ defined by

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

The ***minimum Hamiltonian*** is then defined in the usual manner:

$$H^{\Sigma_{\text{aff}},F_{\mathscr{A}}}_{\min}(\alpha_{v_q}\oplus\beta_{v_q}) = \inf_{u\in U_q} H^{\Sigma_{\text{aff}},F_{\mathscr{A}}}(u,\alpha_{v_q}\oplus\beta_{v_q}).$$

Let $\gamma=(u,c)\in\operatorname{Ctraj}(\Sigma_{\text{aff}},F_{\mathscr{A}})$ with $u$ and $c$ defined on an interval $I$. A LAD one-form field $\lambda\colon[a,b]\to T^*Q$ along $c$ is ***minimising for $(\Sigma_{\text{aff}},F_{\mathscr{A}})$ along $u$*** if

$$H^{\Sigma_{\text{aff}},F_{\mathscr{A}}}(\theta(t)\oplus\lambda(t)) \le H^{\Sigma_{\text{aff}},F_{\mathscr{A}}}_{\min}(\theta(t)\oplus\lambda(t))$$

for almost every $t\in I$ and where

$$\theta(t)=\tfrac{1}{2}T^*(\lambda(t),c'(t))-\nabla_{c'(t)}\lambda(t)-r\lambda_0 f'(A(v_q,\dots,v_q))\hat{A}(c'(t),\dots,c'(t)),\quad t\in[a,b].$$

The result we seek is as follows.

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

*(i) for almost every $t\in[a,b]$ we have*

$$\begin{aligned}
&\nabla^2_{c'(t)}\lambda(t)+R^*(\lambda(t),c'(t))c'(t)-T^*(\nabla_{c'(t)}\lambda(t),c'(t))=\\
&\quad u^a(t)(\nabla Y_a)^*(\lambda(t))+\lambda_0 f'(A(c'(t),\dots,c'(t)))\Big(\nabla A(c'(t),\dots,c'(t))-\\
&\quad r(\nabla_{c'(t)}\hat{A})(c'(t),\dots,c'(t))-r(r-1)u^a(t)\hat{A}(Y_a(c(t)),c'(t),\dots,c'(t))+\\
&\quad rT^*(\hat{A}(c'(t),\dots,c'(t)),c'(t))\Big)-\\
&\quad r\lambda_0 f''(A(c'(t),\dots,c'(t)))\Big(\nabla A(c'(t),\dots,c'(t);c'(t))+\\
&\quad ru^a(t)A(Y_a(c(t)),c'(t),\dots,c'(t))\Big)\hat{A}(c'(t),\dots,c'(t));
\end{aligned}$$

*(ii) $\lambda$ is minimising for $(\Sigma_{\text{aff}},\lambda_0F_{\mathscr{A}})$ along $u$;*

*(iii) either $\lambda_0=1$ or $\theta(a)\oplus\lambda(a)\neq 0$;*

*with*

$$\theta(t)=\tfrac{1}{2}T^*(\lambda(t),c'(t))-\nabla_{c'(t)}\lambda(t)-r\lambda_0 f'(A(v_q,\dots,v_q))\hat{A}(c'(t),\dots,c'(t)),\quad t\in[a,b].$$

*If $U$ is constant then we have*

*(iv) there exists a constant $C\in\mathbb{R}$ so that $H^{\Sigma_{\text{aff}},F_{\mathscr{A}}}(u(t),\theta(t)\oplus\lambda(t))=C$ a.e.*

*If $U$ is constant and if $\gamma=(u,c)$ is a solution of $\mathscr{P}(\Sigma_{\text{aff}},F_{\mathscr{A}},v_{q_0},v_{q_1})$ then condition (iv) is replaced with*

*(v)* $H^{\Sigma_{\mathrm{aff}},F_{\mathscr{A}}}(\theta(t)\oplus\lambda(t))=0$ *a.e.*

*If* $\gamma=(c,u)$ *is a solution of* $\mathscr{P}_{[a,b]}(\Sigma_{\mathrm{aff}},F_{\mathscr{A}},q_0,q_1)$ *then the conditions (i)–(iii) hold and, in addition,* $\lambda(a)=0$ *and* $\lambda(b)=0$*. If* $U$ *is constant then the condition (iv) holds. If* $\gamma$ *is a solution of* $\mathscr{P}(\Sigma_{\mathrm{aff}},F_{\mathscr{A}},q_0,q_1)$ *and if* $U$ *is constant, then condition (v) holds.*

Proof: We will show the equivalence of the conditions in the theorem to those in the general maximum principle stated as Theorem 6.3 in the case when $S_0=\{v_{q_0}\}$ and $S_1=\{v_{q_1}\}$. First we relate the one-form field $\lambda$ along $c$ to the integral curve $\chi$ of the Hamiltonian vector field with Hamiltonian $H^{\Sigma_{\mathrm{aff}},F_{\mathscr{A}}}$ as asserted in the maximum principle. We claim that the curve

$$t\mapsto\Big(\tfrac{1}{2}T^*(\lambda(t),c'(t))-\nabla_{c'(t)}\lambda(t)-r\lambda_0f'(A(v_q,\dots,v_q))\hat{A}(c'(t),\dots,c'(t))\Big)\oplus\lambda(t)\tag{7.2}$$

exactly represents $\chi$ with respect to our splitting of $T^*_{c'(t)}TQ$. To do this we must show that (7.2) is an integral curve of the time-dependent Hamiltonian vector field with Hamiltonian $(t,\alpha_{v_q}\oplus\beta_{v_q})\mapsto H^{\Sigma_{\mathrm{aff}},\lambda_0F_{\mathscr{A}}}(u(t),\alpha_{v_q}\oplus\beta_{v_q})$. Note that $H^{\Sigma_{\mathrm{aff}},\lambda_0F_{\mathscr{A}}}$ is the sum of three functions: (1) $\Lambda_{v_q}\mapsto\Lambda_{v_q}\cdot Z(v_q)$, (2) $\Lambda_{v_q}\mapsto u^a(t)(\Lambda_{v_q}\cdot\mathrm{vlft}(Y_a(q)))$, and (3) $\Lambda_{v_q}\mapsto\lambda_0F_{\mathscr{A}}(u(t),q)$. Thus the Hamiltonian vector field will be the sum of the three Hamiltonian vector fields corresponding to the three Hamiltonians. Let us write these three Hamiltonian vector fields in the splitting of $T_{\Lambda_{v_q}}T^*TQ$. In each case we write, in the usual manner, $\Lambda_{v_q}=\alpha_{v_q}\oplus\beta_{v_q}$. By Propositions 5.8 and 5.1 the Hamiltonian vector field for the Hamiltonian (1) has the representation

$$\alpha_{v_q}\oplus\beta_{v_q}\mapsto v_q\oplus0\oplus\big(R^*(\beta_{v_q},v_q)v_q+\tfrac{1}{2}(\nabla_{v_q}T^*)(\beta_{v_q},v_q)-\tfrac{1}{4}T^*(T^*(\beta_{v_q},v_q),v_q)\big)\oplus(-\alpha_{v_q}).$$

By Proposition 5.13 the Hamiltonian vector field associated with the Hamiltonian (2) has the representation

$$\alpha_{v_q}\oplus\beta_{v_q}\mapsto0\oplus u^a(t)Y_a(q)\oplus\big(u^a(t)\tfrac{1}{2}T^*(\beta_{v_q},Y_a(q))-u^a(t)(\nabla Y_a)^*(\beta_{v_q})\big)\oplus0.$$

Now let us compute the representation of the Hamiltonian vector field associated with the Hamiltonian (3). In this case, we use Lemma 7.3 to see that the Hamiltonian vector field for the Hamiltonian (3) has the representation

$$\begin{aligned}&\alpha_{v_q}\oplus\beta_{v_q}\mapsto\\&\lambda_0f'(A(v_q,\dots,v_q))\big(0\oplus0\oplus\big(-\nabla A(v_q,\dots,v_q)-\tfrac{r}{2}T^*(\hat{A}(v_q\dots,v_q),v_q)\big)\oplus(-r\hat{A}(v_q,\dots,v_q))\big).\end{aligned}$$

Here we have suppressed the explicit dependence of $A$ on $u$, but it should be regarded as being implicit.

We now collect this all together. We write the integral curve of the Hamiltonian vector field as $\theta(t)\oplus\lambda(t)$, similar to Lemma 5.9. From this lemma we then have

$$\begin{aligned}\nabla_{c'(t)}\theta(t)=\;&R^*(\lambda(t),c'(t))c'(t)+\tfrac{1}{2}(\nabla_{c'(t)}T^*)(\lambda(t),c'(t))-\tfrac{1}{4}T^*(T^*(\lambda(t),c'(t)),c'(t))+\\&u^a(t)\tfrac{1}{2}T^*(\lambda(t),Y_a(t))-u^a(t)(\nabla Y_a)^*(\lambda(t))-\\&\lambda_0f'(A(c'(t),\dots,c'(t)))\nabla A(c'(t),\dots,c'(t))-\\&\tfrac{r}{2}\lambda_0f'(A(c'(t),\dots,c'(t)))T^*(\hat{A}(c'(t),\dots,c'(t)),c'(t))+\tfrac{1}{2}T^*(\theta(t),c'(t))\\\nabla_{c'(t)}\lambda(t)=\;&-\theta(t)-r\lambda_0f'(A(c'(t),\dots,c'(t)))\hat{A}(c'(t),\dots,c'(t))+\tfrac{1}{2}T^*(\lambda(t),c'(t)).\end{aligned}$$

Note that the right-hand side of the second equation is LAC since $t \mapsto \theta(t) \oplus \lambda(t)$ is the integral curve for a LIC$^\infty$ Hamiltonian vector field and since $t \mapsto c'(t)$ is LAC. Thus $\lambda$ satisfies a first-order time-dependent differential equation which is LAC in time. Therefore we may conclude that $t \mapsto \lambda(t)$ is LAD. Thus we may covariantly differentiate the second of these equations and substitute the first of the equations into the resulting expression. The result is

$$\begin{aligned}\nabla^2_{c'(t)}\lambda(t) = -R^*(\lambda(t), c'(t))c'(t) + T^*(\nabla_{c'(t)}\lambda(t), c'(t)) + u^a(t)(\nabla Y_a)^*(\lambda(t)) + \\ \lambda_0 f'(A(c'(t), \dots, c'(t)))\Big(\nabla A(c'(t), \dots, c'(t)) - r(\nabla_{c'(t)}\hat{A})(c'(t), \dots, c'(t)) - \\ r(r-1)u^a(t)\hat{A}(Y_a(c(t)), c'(t), \dots, c'(t)) + rT^*(\hat{A}(c'(t), \dots, c'(t)), c'(t))\Big) - \\ r\lambda_0 f''(A(c'(t), \dots, c'(t)))\Big(\nabla A(c'(t), \dots, c'(t); c'(t)) + \\ ru^a(t)A(Y_a(c(t)), c'(t), \dots, c'(t))\Big)\hat{A}(c'(t), \dots, c'(t))\end{aligned}$$

which holds a.e. From this we conclude that $t \mapsto \chi(t)$ as defined by (7.2) has the property that $\chi'(t) = X_{H_u^{\Sigma_{\mathrm{aff}}, \lambda_0 F_{\mathscr{A}}}}$ a.e. Thus the existence of $\lambda$ satisfying (i) is equivalent to the existence of the integral curve $\chi$ as asserted in Theorem 6.3.

We note that for $u \in \mathbb{R}^m$, the vector field $v_q \mapsto Z(v_q) + u^a Y_a(q)$ on $TQ$ has the form $v_q \mapsto v_q \oplus u^a Y_a(q)$ in the splitting of $TTQ$ defined in Section 5.7. This shows that the Hamiltonian $H^{\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}}$ has the given form in the splitting of $T^*TQ$.

It is then clear that the conditions (ii)–(v) are equivalent to the conditions (iii)–(vi) of Theorem 6.3.

The final assertions regarding solutions of $\mathscr{P}_{[a,b]}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1)$ and $\mathscr{P}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1)$ follow from Theorem 6.3 in the case where $S_0 = T_{q_0}Q$ and $S_1 = T_{q_1}Q$. Note that in the splitting $T^*_{v_{q_0}}TQ = T^*_{q_0}Q \oplus T^*_{q_0}Q$, we have $\mathrm{ann}(T_v(T_{q_0}Q)) = T^*_{q_0}Q \oplus \{0\}$ for $v \in T_{q_0}Q$, and similarly for $q_1$. ■

**7.8 Remarks:** 1. Let us consider the import of the preceding theorem. Were we to simply apply the maximum principle of Theorem 6.3, we would obtain a first-order differential equation for a one-form field along trajectories in $TQ$. Theorem 7.7 provides a second-order differential equation for a one-form field along trajectories in $Q$. But, more importantly, the differential equation governing the evolution of this one-form field on $Q$ provides a clear indication of how the geometry of the control system enters into the optimal control problem. We shall exploit this knowledge in the next section to formulate an optimal control problem which clearly utilises the geometry of the control system through its affine connection.

2. We call the second-order equation for $\lambda$ in (i) of Theorem 7.7 the ***adjoint equation*** for $(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}})$. Note that the left-hand side of this equation is none other than the adjoint Jacobi equation. One of the features of our constructions here is that we are able to intrinsically provide an equation for the adjoint covector field which is decoupled from the control system equations. This is generally not possible when talking about control systems on manifolds, but is possible here because of the existence of the myriad Ehresmann connections attached to the affine connection $\nabla$.

3. It is an interesting and useful consequence of our use of an input range map that we may trivially incorporate systems with potential energy. Let us describe how this may be done. One has a potential function $V$ on $Q$ and defines a Lagrangian by $L(v_q) = \frac{1}{2}g(v_q, v_q) - V(q)$ where $g$ is a Riemannian metric on $Q$. The equations of motion for the controlled system with control vector fields $\mathscr{Y} = \{Y_1, \dots, Y_m\}$ are then

$$\overset{g}{\nabla}_{c'(t)}c'(t) = -\operatorname{grad} V(c(t)) + u^a(t)Y_a(c(t))$$

where $\operatorname{grad} V$ is the gradient vector field associated with $V$. Clearly it makes no difference in the general scheme if we replace $\overset{g}{\nabla}$ with an arbitrary affine connection $\nabla$, and replace $\operatorname{grad} V$ with a general vector field $Y_0$ on $Q$. The question is how to incorporate the vector field $Y_0$ into our affine connection control system $(Q, \nabla, \mathscr{Y}, U)$. We do this by defining a new set $\tilde{\mathscr{Y}} = \{Y_0, Y_1, \dots, Y_m\}$ of control vector fields and a new input range map $\tilde{U}_q = \{1\} \times U_q \subset \mathbb{R} \times \mathbb{R}^m \simeq \mathbb{R}^{m+1}$. One may now apply verbatim Theorem 7.7 to the new affine connection control system $\tilde{\Sigma}_{\text{aff}} = (Q, \nabla, \tilde{\mathscr{Y}}, \tilde{U})$. •

Let us introduce some terminology which we can use to talk about controlled arcs which satisfy the conditions of Theorem 7.7.

**7.9 Definition:** Let $\Sigma_{\text{aff}}$ be an affine connection control system and let $F_{\mathscr{A}}$ be a cost function for $\Sigma_{\text{aff}}$. Let $\mathscr{P}$ be one of the problems $\mathscr{P}(\Sigma_{\text{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1})$, $\mathscr{P}_{[a,b]}(\Sigma_{\text{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1})$, $\mathscr{P}(\Sigma_{\text{aff}}, F_{\mathscr{A}}, q_0, q_1)$, or $\mathscr{P}_{[a,b]}(\Sigma_{\text{aff}}, F_{\mathscr{A}}, q_0, q_1)$, and let $\gamma = (c, u) \in \operatorname{Carc}(\Sigma_{\text{aff}}, F_{\mathscr{A}})$ satisfy the necessary conditions of Theorem 7.7 with $\lambda_0 \in \{0, 1\}$ and $\lambda$ as given by the theorem. We then call $\gamma$ a ***controlled extremal*** for $\mathscr{P}$, $\lambda_0$ the ***constant Lagrange multiplier*** for $\gamma$, and $\lambda$ the ***adjoint vector field*** for $\gamma$. If $\lambda_0 = 1$ (resp. $\lambda_0 = 0$) then $\gamma$ is a ***normal*** (resp. ***abnormal***) controlled extremal. •

The version of the maximum principle stated in Theorem 7.7 is quite general, and much of the complexity in its statement is owed to that generality. However, when looking at specific classes of optimal problems, the general form of the theorem can often be reduced to something more appealing, and we next demonstrate this by beginning to look at a simple class of cost function.

## 8. Force minimising controls

With the backdrop of our general discussion of Section 7, let us now move on to consider an interesting special case of an optimal control problem involving minimising a function of the inputs. Since the inputs are coefficients of the input vector fields, and the input vector fields are related to forces in physical systems, we dub this the force minimisation problem. We suppose that $Q$ is equipped with a Riemannian metric $g$. The cost function we consider is

$$F(u, v_q) = \frac{1}{2}g(u^aY_a(q), u^bY_b(q)). \tag{8.1}$$

In the parlance of Section 7.2, we use a $(0, 0)$ $\mathbb{R}^m$-dependent tensor field and we choose $f = \mathrm{id}_{\mathbb{R}}$. We choose for our input range map $U\colon Q \to \mathbf{2}^{\mathbb{R}^m}$ the map defined by $U_q = \mathbb{R}^m$ for each $q \in Q$.

For the sake of formality, let us define precisely the problem we are solving.

**8.1 Definition:** Let $\Sigma_{\text{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be an affine connection control system with $U_q = \mathbb{R}^m$ for $q \in Q$ and with cost function $F$ as defined by (8.1). Let $q_0, q_1 \in Q$, and let $v_{q_0} \in T_{q_0}Q$ and $v_{q_1} \in T_{q_1}Q$.

(i) A controlled arc $\gamma = (u, c)$ is a ***solution of*** $\boldsymbol{\mathscr{F}(\Sigma_{\text{aff}}, v_{q_0}, v_{q_1})}$ if it is a solution of $\mathscr{P}(\Sigma_{\text{aff}}, F, v_{q_0}, v_{q_1})$.

(ii) A controlled arc $\gamma = (u, c)$ is a ***solution of*** $\boldsymbol{\mathscr{F}_{[a,b]}(\Sigma_{\text{aff}}, v_{q_0}, v_{q_1})}$ if it is a solution of $\mathscr{P}_{[a,b]}(\Sigma_{\text{aff}}, F, v_{q_0}, v_{q_1})$.

(iii) A controlled arc $\gamma = (u, c)$ is a ***solution of*** $\boldsymbol{\mathscr{F}(\Sigma_{\text{aff}}, q_0, q_1)}$ if it is a solution of $\mathscr{P}(\Sigma_{\text{aff}}, F, q_0, q_1)$.

(iv) A controlled arc $\gamma = (u, c)$ is a ***solution of*** $\boldsymbol{\mathscr{F}_{[a,b]}(\Sigma_{\text{aff}}, q_0, q_1)}$ if it is a solution of $\mathscr{P}_{[a,b]}(\Sigma_{\text{aff}}, F, q_0, q_1)$. •

It will be helpful to make a few straightforward constructions given the data for the force minimisation optimal control problems. We define a distribution $\mathsf{Y}$ on $Q$ by

$$\mathsf{Y}_q = \operatorname{span}_{\mathbb{R}}(Y_1(q), \dots, Y_m(q)),$$

and we suppose this distribution to have constant rank (but not necessarily rank $m$). The map $i_{\mathsf{Y}} \colon \mathsf{Y} \to TQ$ denotes the inclusion. The map $P_{\mathsf{Y}} \colon TQ \to TQ$ denotes the $g$-orthogonal projection onto the distribution $\mathsf{Y}$, with $P_{\mathsf{Y}_q}$ being its restriction to $T_qQ$. We may then define a $(0,2)$ tensor field $g_{\mathsf{Y}}$ on $Q$ by $g_{\mathsf{Y}}|T_qQ = P^*_{\mathsf{Y}_q}(g|T_qQ)$. That is, $g_{\mathsf{Y}}$ is the restriction to $\mathsf{Y}$ of $g$. We also have the associated $(2,0)$ tensor $h_{\mathsf{Y}}$ defined by

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

Here $g^\sharp \colon T^*Q \to TQ$ is the canonical musical isomorphism associated with the Riemannian metric $g$. We also define the vector bundle mapping $h^\sharp_{\mathsf{Y}} \colon T^*Q \to TQ$ by $\langle \alpha_q; h^\sharp_{\mathsf{Y}}(\beta_q)\rangle = h_{\mathsf{Y}}(\alpha_q, \beta_q)$.

**8.1. General affine connections.** We first look at the case when $\Sigma_{\text{aff}} = (Q, \nabla, \mathscr{Y}, U)$ is a general affine connection control system with $U$ as specified above. Thus, in particular $\nabla$ is not the Levi-Civita connection associated with the Riemannian metric $g$ used to define the cost function.

The Hamiltonian function on $\mathscr{D}^*_{TQ} = \mathbb{R}^m \times T^*TQ$ is given by

$$H^{\Sigma_{\text{aff}},F}(u, \alpha_{v_q} \oplus \beta_{v_q}) = \frac{1}{2} g(u^a Y_a(q), u^b Y_b(q)) + \alpha_{v_q} \cdot v_q + u^a(\beta_{v_q} \cdot Y_a(q)).$$

Let us define $A_{\text{sing}}(\Sigma_{\text{aff}}) \subset T^*TQ$ by

$$A_{\text{sing}}(\Sigma_{\text{aff}}) = \{\alpha_{v_q} \oplus \beta_{v_q} \mid \beta_{v_q} \in \operatorname{ann}(\mathsf{Y}_q\}).$$

Thus the restriction of $H^{\Sigma_{\text{aff}},F}$ to $\mathbb{R}^m \times A_{\text{sing}}(\Sigma_{\text{aff}})$ is independent of $u$.

The following result gives the form of the minimum Hamiltonian and the values of $u$ by which the minima are realised.

**8.2 Lemma:** *The following statements hold.*

*(i) The minimum Hamiltonian for the cost function $F$ is given by*

$$H^{\Sigma_{\text{aff}},F}_{\min}(\alpha_{v_q} \oplus \beta_{v_q}) = \alpha_{v_q} \cdot v_q - \tfrac{1}{2} h_{\mathsf{Y}}(\beta_{v_q}, \beta_{v_q}).$$

*If $u \in \mathbb{R}^m$ is a point at which $H^{\Sigma_{\text{aff}},F}_{\min}$ is realised, then $u$ is determined by*

$$u^a Y_a(q) = -P_{\mathsf{Y}_q}(g^\sharp(\beta_{v_q})). \tag{8.2}$$

*(ii) The minimum Hamiltonian with zero cost function is*

$$H^{\Sigma_{\text{aff}},0}_{\min}(\alpha_{v_q} \oplus \beta_{v_q}) = \begin{cases} \alpha_{v_q} \cdot v_q, & \alpha_{v_q} \oplus \beta_{v_q} \in A_{\text{sing}}(\Sigma_{\text{aff}}) \\ -\infty, & \textit{otherwise.} \end{cases}$$

*The values of $u \in \mathbb{R}^m$ which realise $H^{\Sigma_{\text{aff}},0}_{\min}$ at $\alpha_{v_q} \oplus \beta_{v_q} \in A_{\text{sing}}(\Sigma_{\text{aff}})$ are arbitrary.*

Proof: (i) We fix the state $\alpha_{v_q} \oplus \beta_{v_q}$ and determine $u$ so as to minimise $H^{\Sigma_{\text{aff}},F}(\alpha_{v_q} \oplus \beta_{v_q})$. We first note that, with the state fixed, we may think of $H^{\Sigma_{\text{aff}},F}$ as being a function on the subspace $\mathsf{Y}_q$ of $T_qQ$. Let us denote a typical point in $\mathsf{Y}_q$ by $w$ and note that $H^{\Sigma_{\text{aff}},F}$ as a function on $\mathsf{Y}_q$ is

$$\begin{aligned} w \mapsto\ & \tfrac{1}{2} g(i_{\mathsf{Y}_q}(w), i_{\mathsf{Y}_q}(w)) + \alpha_{v_q} \cdot v_q + \beta_{v_q} \cdot i_{\mathsf{Y}_q}(w) \\ &= \tfrac{1}{2} g(i_{\mathsf{Y}_q}(w), i_{\mathsf{Y}_q}(w)) + \alpha_{v_q} \cdot v_q + g(g^\sharp(\beta_{v_q}), i_{\mathsf{Y}_q}(w)) \\ &= \tfrac{1}{2} g(i_{\mathsf{Y}_q}(w), i_{\mathsf{Y}_q}(w)) + \alpha_{v_q} \cdot v_q + g(P_{\mathsf{Y}_q}(g^\sharp(\beta_{v_q})), i_{\mathsf{Y}_q}(w)). \end{aligned}$$

Since this is a convex function of $w$, it will have a unique minimum. Differentiating $H^{\Sigma_{\text{aff}},F}$ with respect to $w$ and setting the resulting expression to 0 shows that the minimum satisfies

$$i_{\mathsf{Y}}(w_{\min}) = -P_{\mathsf{Y}_q}(g^\sharp(\beta_{v_q})).$$

Controls $u$ which give $w_{\min}$ are thus as specified by (8.2). This part of the lemma is then proved by substituting the expression for $w_{\min}$ into $H^{\Sigma_{\text{aff}},F}$.

(ii) For zero cost function, $H^{\Sigma_{\text{aff}},0}$ is an affine function of $u$. Thus it will be bounded below if and only if the linear part is zero. This happens if and only if $\alpha_{v_q} \oplus \beta_{v_q} \in A_{\text{sing}}(\Sigma_{\text{aff}})$. In this case $H^{\Sigma_{\text{aff}},0}$ is a constant function of $u$ and so $u$ is left unspecified by asking that it realise $H^{\Sigma_{\text{aff}},0}_{\min}$ at those points in $A_{\text{sing}}(\Sigma_{\text{aff}})$. ■

**8.3 Remark:** Note that the controls are uniquely determined by the state if the vector fields $Y_1, \dots, Y_m$ are linearly independent. Otherwise, there will be multiple vectors $u$ which satisfy (8.2), all of which give rise to the same minimum Hamiltonian $H^{\Sigma_{\text{aff}},F}_{\min}$. Note that in using Theorem 7.7 we allow $\lambda_0 = 0$ for and only for initial conditions lying in $A_{\text{sing}}(\Sigma_{\text{aff}})$. •

It is possible to determine a simplified form for the controlled extremals for the force minimisation problem. Let $\mathrm{L}(\mathrm{ann}(\mathsf{Y}) \times \mathsf{Y}; T^*Q)$ denote the vector bundle of multilinear bundle maps from $\mathrm{ann}(\mathsf{Y}) \times \mathsf{Y}$ to $T^*Q$. To describe the abnormal extremals it will be helpful to define $B_{\mathsf{Y}} \in \Gamma^\infty(\mathrm{L}(\mathrm{ann}(\mathsf{Y}) \times \mathsf{Y}; T^*Q))$ by

$$\langle B_{\mathsf{Y}}(\alpha, Y); X \rangle = \langle \alpha; \nabla_X Y \rangle.$$

That $B_{\mathsf{Y}}(\alpha, Y)$ does not depend on the derivative of $Y$ follows since for a function $f \in C^\infty(Q)$ we compute

$$\begin{aligned}\langle B_{\mathsf{Y}}(\alpha, fY); X\rangle &= \langle \alpha; \nabla_X(fY)\rangle \\ &= \langle \alpha; f\nabla_X Y\rangle + \langle \alpha; (\mathscr{L}_X f)Y\rangle \\ &= \langle f B_{\mathsf{Y}}(\alpha, Y); X\rangle,\end{aligned}$$

since $Y \in \mathsf{Y}$ and $\lambda \in \operatorname{ann}(\mathsf{Y})$. We may now state the following theorem.

**8.4 Theorem:** *Let $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be an affine connection control system with $U_q = \mathbb{R}^m$ for $q \in Q$. Suppose that $\gamma = (u, c)$ is a controlled extremal for $\mathscr{F}_{[a,b]}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1})$ or for $\mathscr{F}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1})$ with $u$ and $c$ defined on $[a, b]$, and with $\lambda$ the adjoint vector field and $\lambda_0$ the Lagrange multiplier. We have the following two situations.*

*(i) $\lambda_0 = 1$: In this case it is necessary and sufficient that $c$ and $\lambda$ together satisfy the differential equations*

$$\begin{aligned}&\nabla_{c'(t)} c'(t) = -h^\sharp_{\mathsf{Y}}(\lambda(t)) \\ &\nabla^2_{c'(t)}\lambda(t) + R^*(\lambda(t), c'(t))c'(t) - T^*(\nabla_{c'(t)}\lambda(t), c'(t)) = \tfrac{1}{2}\nabla h_{\mathsf{Y}}(\lambda(t), \lambda(t)) - \\ &\qquad T^*(\lambda(t), h^\sharp_{\mathsf{Y}}(\lambda(t))).\end{aligned} \tag{8.3}$$

*(ii) $\lambda_0 = 0$: In this case it is necessary and sufficient that*

*(a) $\nabla_{c'(t)} c'(t) = u^a(t) Y_a(c(t))$,*

*(b) $\lambda(t) \in \operatorname{ann}(\mathsf{Y}_{c(t)})$ for $t \in [a, b]$ and*

*(c) $\lambda$ satisfies the equation along $c$ given by:*

$$\nabla^2_{c'(t)}\lambda(t) + R^*(\lambda(t), c'(t))c'(t) - T^*(\nabla_{c'(t)}\lambda(t), c'(t)) = B_{\mathsf{Y}}(\lambda(t), u^a(t)Y_a(t)).$$

*If $\gamma = (u, c)$ is a solution of $\mathscr{F}_{[a,b]}(\Sigma_{\mathrm{aff}}, q_0, q_1)$ or of $\mathscr{F}(\Sigma_{\mathrm{aff}}, q_0, q_1)$ then we additionally have $\lambda(a) = 0$ and $\lambda(b) = 0$.*

Proof: (i) We note that Lemma 8.2 provides for us a minimum Hamiltonian which is smooth. Since the input range map $U$ is constant we may conclude that $H^{\Sigma_{\mathrm{aff}},F}_{\min}$ is constant a.e. along the solutions. We conclude that to compute the extremal arcs is suffices to compute trajectories of the Hamiltonian vector field with Hamiltonian $H^{\Sigma_{\mathrm{aff}},F}_{\min}$. We thus simply compute the equations corresponding to the Hamiltonian vector field with Hamiltonian

$$H^{\Sigma_{\mathrm{aff}},F}_{\min}(\alpha_{v_q} \oplus \beta_{v_q}) = \alpha_{v_q} \cdot v_q - \frac{1}{2} h_{\mathsf{Y}}(\beta_{v_q}, \beta_{v_q}).$$

As usual we use the notation corresponding to the splitting of $T^*TQ$. We also write the vector field in the splitting of $T(T^*TQ)$ as we have been doing all along. Thus the Hamiltonian vector field with Hamiltonian $H^{\Sigma_{\mathrm{aff}},F}_{\min}$ is the sum of two Hamiltonians. By Theorem 5.10 the first Hamiltonian vector field has the representation

$$v_q \oplus 0 \oplus \big(R^*(\beta_{v_q}, v_q)v_q + \tfrac{1}{2}(\nabla_{v_q}T^*)(\beta_{v_q}, v_q) - \tfrac{1}{4}T^*(T^*(\beta_{v_q}, v_q), v_q)\big) \oplus (-\alpha_{v_q}).$$

By Lemma 7.4 the Hamiltonian vector field for the Hamiltonian $-\frac{1}{2}h_{\mathsf{Y}}(\beta_{v_q}, \beta_{v_q})$ is

$$0 \oplus (-h^\sharp_{\mathsf{Y}}(\beta_{v_q})) \oplus \big(-\tfrac{1}{2}\nabla h_{\mathsf{Y}}(\beta_{v_q}, \beta_{v_q}) + \tfrac{1}{2}T^*(\beta_{v_q}, h^\sharp_{\mathsf{Y}}(\beta_{v_q}))\big) \oplus 0.$$

This immediately gives

$$\nabla_{c'(t)}c'(t) = -h^\sharp_{\mathsf{Y}}(\lambda(t))$$

which is the first of equations (8.3). Now let $\theta(t) \oplus \lambda(t)$ be the integral curve over $c'$ of the Hamiltonian vector field. By Lemma 5.9 we have

$$\begin{aligned}\nabla_{c'(t)}\theta(t) = {} & R^*(\lambda(t), c'(t))c'(t) + \tfrac{1}{2}(\nabla_{c'(t)}T^*)(\lambda(t), c'(t)) - \tfrac{1}{4}T^*(T^*(\lambda(t), c'(t)))- \\ & \tfrac{1}{2}\nabla h_{\mathsf{Y}}(\lambda(t), \lambda(t)) + \tfrac{1}{2}T^*(\lambda(t), h^\sharp_{\mathsf{Y}}(\lambda(t))) + \tfrac{1}{2}T^*(\theta(t), c'(t)) \\ \nabla_{c'(t)}\lambda(t) = {} & -\theta(t) + \tfrac{1}{2}T^*(\lambda(t), c'(t)).\end{aligned} \tag{8.4}$$

Covariantly differentiating the second equation along $c$ gives

$$\begin{aligned}\nabla^2_{c'(t)}\lambda(t) = -\nabla_{c'(t)}\theta(t) + \tfrac{1}{2}(\nabla_{c'(t)}T^*)(\lambda(t), c'(t)) + \\ \tfrac{1}{2}T^*(\nabla_{c'(t)}\lambda(t), c'(t)) - \tfrac{1}{2}T^*(\lambda(t), h^\sharp_{\mathsf{Y}}(\lambda(t))).\end{aligned}$$

Substituting the first of equations (8.4) gives the second of equations (8.3) which thus completes the proof of this part of the lemma.

(ii) We first note that $\lambda$ can be minimising only if $\theta(t) \oplus \lambda(t) \in A_{\text{sing}}(\Sigma_{\text{aff}})$ for all $t \in [a, b]$. This means that $\lambda(t)$ must annihilate $\mathsf{Y}_{c(t)}$. Since $\lambda_0 = 0$, our result follows from Theorem 7.7 and the definition of $B_{\mathsf{Y}}$. ■

**8.5 Remarks:** 1. The theorem implies that all normal extremals for the force minimisation problem are of class $C^\infty$.

2. If one happens to choose $v_{q_0}, v_{q_1} \in TQ$ with the property that there is a geodesic $c\colon [0, T] \to Q$ satisfying $c'(0) = v_{q_0}$ and $c'(T) = v_{q_1}$, then the optimal control for the problems $\mathscr{F}(\Sigma_{\text{aff}}, v_{q_0}, v_{q_1})$ and $\mathscr{F}_{[a,a+T]}(\Sigma_{\text{aff}}, v_{q_0}, v_{q_1})$, $a \in \mathbb{R}$, is the zero control. The associated extremal is readily seen to be both normal and abnormal.

3. The matter of investigating the existence of strictly abnormal extremals (i.e., abnormal extremals which are not simultaneously normal) which are also minimisers would appear likely to take on a flavour similar to that of the sub-Riemannian case [see Liu and Sussmann 1994, Montgomery 1994]. This will be the subject of future work.

4. If $\gamma = (u, c) \in \text{Ctraj}(\Sigma_{\text{aff}})$ then

$$\begin{aligned}\nabla_{c'(t)}c'(t) = u^a(t)Y_a(c(t)) \quad \Longrightarrow \\ g(u^a(t)Y_a(c(t)), u^b(t)Y_b(c(t))) = g(\nabla_{c'(t)}c'(t), \nabla_{c'(t)}c'(t)).\end{aligned}$$

Thus the force minimisation problem may be seen as our wishing to minimise

$$\int_a^b g(\nabla_{c'(t)}c'(t), \nabla_{c'(t)}c'(t))\,\mathrm{d}t$$

over all LAD curves $c\colon [a, b] \to Q$ subject to certain boundary conditions (fixed or free velocity) and subject to the constraint that $\nabla_{c'(t)}c'(t) \in \mathsf{Y}_{c(t)}$ a.e. This may be thought of as a higher-order version of the sub-Riemannian geodesic problem. Indeed note that the equations (8.3) for the controlled extremals involve only the restriction of $g$ to the distribution $\mathsf{Y}$. In the fully actuated case (see next section) we have a classical calculus of variations problem with a Lagrangian depending on first and second time-derivatives. This is the approach of Crouch and Silva Leite [1991] and Noakes, Heinzinger, and Paden [1989]. •

**8.2. The fully actuated case.** As mentioned in the introduction, Crouch and Silva Leite [1991] and Noakes, Heinzinger, and Paden [1989] consider the force minimisation problem with the Levi-Civita connection and with full actuation. Let us now consider the *general* case with full actuation. First let's be precise about the meaning of full actuation. An affine connection control system $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ is ***fully actuated*** if $T_qQ = \operatorname{span}_{\mathbb{R}}(Y_1(q), \dots, Y_m(q))$ for each $q \in Q$. Thus in this section we let $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be a fully actuated affine connection control system where $U_q = \mathbb{R}^m$ for each $q \in Q$.

Let us first show that all extremals for the fully actuated force minimisation problem are normal.

**8.6 Proposition:** *Let $\Sigma_{\mathrm{aff}} = (Q, \nabla, \mathscr{Y}, U)$ be a fully actuated affine connection control system with $\gamma = (u, c)$ a controlled extremal for one of the four problems of Definition 8.1. The corresponding Lagrange multiplier $\lambda_0$ is nonzero.*

Proof: This follows from the Hamiltonian minimisation condition. Since the Hamiltonian is

$$H^{\Sigma_{\mathrm{aff}},\lambda_0 F}(u, \alpha_{v_q} \oplus \beta_{v_q}) = \frac{\lambda_0}{2} g(u^a Y_a(q), u^b Y_b(q)) + \alpha_{v_q} \cdot v_q + u^a(\beta_{v_q} \cdot Y_a(q)),$$

the only way for the Hamiltonian to be minimum with $\lambda_0 = 0$ is for $\beta_{v_q}$ to be zero. This cannot happen since Theorem 7.7 asserts that both $\lambda_0$ and $\lambda$ cannot be zero along an extremal. ■

We may now concentrate on the normal case of Theorem 8.4. The simplification here arises since $h_\gamma$ becomes the vector bundle metric $g^{-1}$ on $T^*Q$ induced by $g$. From Theorem 8.4, if $\gamma = (u, c)$ is a solution of $\mathscr{F}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1})$ we have

$$\begin{aligned} &\nabla_{c'(t)} c'(t) = -g^\sharp(\lambda(t)) \\ &\nabla^2_{c'(t)} \lambda(t) + R^*(\lambda(t), c'(t))c'(t) - T^*(\nabla_{c'(t)}\lambda(t), c'(t)) = \tfrac{1}{2}\nabla g^{-1}(\lambda(t), \lambda(t)) - \\ &\qquad T^*(\lambda(t), g^\sharp(\lambda(t))). \end{aligned} \tag{8.5}$$

We immediately see that the adjoint vector field $\lambda$ is determined algebraically from the covariant derivative of $c$ along itself. This allows us to eliminate the adjoint vector field from the equations (8.5) as the following result asserts.

**8.7 Proposition:** *Let $\Sigma_{\mathrm{aff}}$ be a fully actuated affine connection control system, let $\gamma = (u, c)$ be a controlled extremal for one of the four control problems of Definition 8.1 with $u$ and $c$ defined on $[a, b]$, and let $\lambda\colon [a, b] \to T^*Q$ be the corresponding adjoint vector field. Then $\lambda(t) = -g^\flat(\nabla_{c'(t)} c'(t))$ for $t \in [a, b]$ and $c$ satisfies the equation*

$$\begin{aligned} &\nabla^3_{c'(t)} c'(t) + g^\sharp(R^*(g^\flat(\nabla_{c'(t)} c'(t)), c'(t))c'(t) + g^\sharp(T^*(g^\flat(\nabla_{c'(t)} c'(t)), c'(t))) + \\ &\quad g^\sharp(T^*(g^\flat((\nabla_{c'(t)} g^\sharp)(g^\flat(\nabla_{c'(t)} c'(t)))))) - \tfrac{1}{2} g^\sharp(\nabla g(g^\flat(\nabla_{c'(t)} c'(t)), g^\flat(\nabla_{c'(t)} c'(t)))) - \\ &\qquad g^\sharp(T^*(g^\flat(\nabla_{c'(t)} c'(t)), \nabla_{c'(t)} c'(t))) + (\nabla^2_{c'(t)} g^\sharp)(g^\flat(\nabla_{c'(t)} c'(t))) - \\ &\qquad 2(\nabla_{c'(t)} g^\sharp)(g^\flat(\nabla_{c'(t)} c'(t))) - 2(\nabla_{c'(t)} g^\sharp)(g^\flat((\nabla_{c'(t)} g^\sharp)(g^\flat(\nabla_{c'(t)} c'(t))))) + \\ &\qquad\qquad g^\sharp(\nabla^2_{c'(t)} \lambda(t)) = 0. \end{aligned}$$

*If $\gamma = (u, c)$ is a solution of either $\mathscr{F}(\Sigma_{\mathrm{aff}}, q_0, q_1)$ or $\mathscr{F}_{[a,b]}(\Sigma_{\mathrm{aff}}, q_0, q_1)$ then we additionally must have $\nabla_{c'(a)} c'(a) = 0$ and $\nabla_{c'(b)} c'(b) = 0$.*

Proof: If we covariantly differentiate the first of equations (8.3) along $c$ we get

$$\nabla^2_{c'(t)}c'(t) = -(\nabla_{c'(t)}g^\sharp)(\lambda(t)) - g^\sharp(\nabla_{c'(t)}\lambda(t))$$

and differentiating the same way again gives

$$\begin{aligned}\nabla^3_{c'(t)}c'(t) = &-(\nabla^2_{c'(t)}g^\sharp)(\lambda(t)) - 2(\nabla_{c'(t)}g^\sharp)(\nabla_{c'(t)}\lambda(t)) - g^\sharp(\nabla^2_{c'(t)}\lambda(t))\\ = &-(\nabla^2_{c'(t)}g^\sharp)(g^\flat(\nabla_{c'(t)}c'(t))) + 2(\nabla_{c'(t)}g^\sharp)(g^\flat(\nabla_{c'(t)}c'(t)))+\\ &2(\nabla_{c'(t)}g^\sharp)(g^\flat((\nabla_{c'(t)}g^\sharp)(g^\flat(\nabla_{c'(t)}c'(t))))) - g^\sharp(\nabla^2_{c'(t)}\lambda(t)).\end{aligned}$$

If we now employ the second of equations (8.3) the result follows. ■

**8.3. The Levi-Civita affine connection.** Now we specialise the constructions of the previous sections to the case when $\nabla = \overset{g}{\nabla}$, the Levi-Civita connection determined by the Riemannian metric used in the definition of the cost function. In this case, matters simplify somewhat since $\overset{g}{\nabla} g = 0$ and since $\overset{g}{\nabla}$ is torsion-free. Let us state the result Theorem 8.4 for Levi-Civita connections.

**8.8 Proposition:** *Let $g$ be a Riemannian metric on $Q$ and consider the affine connection control system $\Sigma_{\text{aff}} = (Q, \overset{g}{\nabla}, \mathscr{Y}, U)$ with cost function $F$ defined using $g$. Suppose that $\gamma = (u, c)$ is a solution of one of the four control problems of Definition 8.1 with $u$ and $c$ defined on $[a, b]$. Then $c$ is of class $C^\infty$ and there exists a $C^\infty$ vector field $w$ along $c$ so that together $c$ and $w$ satisfy the equations*

$$\begin{aligned}&\overset{g}{\nabla}_{c'(t)}c'(t) = -P_{\mathsf{Y}}(w(t))\\ &\overset{g}{\nabla}{}^2_{c'(t)}w(t) + \overset{g}{R}(w(t), c'(t))c'(t) = \tfrac{1}{2}g^\sharp(\overset{g}{\nabla}g_{\mathsf{Y}}(w(t), w(t))).\end{aligned} \tag{8.6}$$

*If $\gamma = (u, c)$ is a solution of $\mathscr{F}(\Sigma_{\text{aff}}, q_0, q_1)$ or $\mathscr{F}_{[a,b]}(\Sigma_{\text{aff}}, q_0, q_1)$ then we additionally have $w(a) = 0$ and $w(b) = 0$.*

Proof: Let $\lambda$ be the one-form field along $c$ as given by Theorem 8.4, and define $w = g^\sharp \circ \lambda$. We shall show that $c$ and $w$ satisfy equations (8.6). Since $\overset{g}{\nabla} g = 0$ we have

$$\overset{g}{\nabla}{}^2_{c'(t)}w(t) = g^\sharp(\overset{g}{\nabla}{}^2_{c'(t)}\lambda(t)). \tag{8.7}$$

By equation (5.17) from the proof of Proposition 5.12 we have

$$\overset{g}{R}(w(t), c'(t))c'(t) = g^\sharp(\overset{g}{R}{}^*(\lambda(t), c'(t))c'(t)). \tag{8.8}$$

Now let $\beta \in \Gamma^\infty(T^*Q)$ and $X = g^\sharp(\beta) \in \Gamma^\infty(TQ)$. From the definition of $h_{\mathsf{Y}}$ we have

$$h_{\mathsf{Y}}(\beta(q), \beta(q)) = g_{\mathsf{Y}}(X(q), X(q)).$$

Therefore, for $u \in T_qQ$ we have

$$\overset{g}{\nabla}_u h_{\mathsf{Y}}(\beta(q), \beta(q)) + 2h_{\mathsf{Y}}(\overset{g}{\nabla}_u\beta(q), \beta(q)) = \overset{g}{\nabla}_u g_{\mathsf{Y}}(X(q), X(q)) + 2g_{\mathsf{Y}}(\overset{g}{\nabla}_u X(q), X(q)).$$

Since $\overset{g}{\nabla} g = 0$ we have $\nabla_u \beta(q) = g^\sharp(\nabla_u X(q))$ from which we ascertain that

$$\overset{g}{\nabla}_u h_{\mathsf{Y}}(\lambda(t), \lambda(t)) = \overset{g}{\nabla}_u g_{\mathsf{Y}}(w(t), w(t)). \tag{8.9}$$

Bringing together equations (8.7), (8.8), and (8.9) and the definition of $h_{\mathsf{Y}}$ gives the result by virtue of Theorem 8.4. ■

Now let us specialise to the fully actuated case. One applies Proposition 8.7 to show that the fully actuated controlled extremals satisfy

$$\overset{g}{\nabla}{}^3_{c'(t)} c'(t) + g^\sharp\big(\overset{g}{R}{}^*(g^\flat(\overset{g}{\nabla}_{c'(t)} c'(t)), c'(t))c'(t)\big) = 0.$$

Now we recall the equation (5.17) from the proof of Proposition 5.12 to prove the following result which agrees with Crouch and Silva Leite [1991] and Noakes, Heinzinger, and Paden [1989].

**8.9 Proposition:** *Let $g$ be a Riemannian metric on $Q$ and consider the fully actuated affine connection control system $\Sigma_{\text{aff}} = (Q, \overset{g}{\nabla}, \mathscr{Y}, U)$ with cost function $F$ defined using $g$. A controlled extremal $\gamma = (u, c)$ for one of the four control problems of Definition 8.1 with $u$ and $c$ defined on $[a, b]$ satisfies the differential equation*

$$\overset{g}{\nabla}{}^3_{c'(t)} c'(t) + \overset{g}{R}(\overset{g}{\nabla}_{c'(t)} c'(t), c'(t))c'(t) = 0.$$

*If $\gamma = (u, c)$ is either a solution of $\mathscr{F}(\Sigma_{\text{aff}}, q_0, q_1)$ or $\mathscr{F}_{[a,b]}(\Sigma_{\text{aff}}, q_0, q_1)$ then we additionally have $\overset{g}{\nabla}_{c'(a)} c'(a) = 0$ and $\overset{g}{\nabla}_{c'(b)} c'(b) = 0$.*

## List of symbols and abbreviations

| | |
|---|---|
| $\langle \cdot ; \cdot \rangle$ | : the natural pairing of a vector space with its dual, 5 |
| $g^\flat$ | : the vector bundle map from $TQ$ to $T^*Q$ associated with a symmetric $(0, 2)$ tensor $g$, 27 |
| $h^\sharp$ | : the vector bundle map from $T^*Q$ to $TQ$ associated with a symmetric $(2, 0)$ tensor $h$, 27 |
| $\mathbf{2}^S$ | : the set of subsets of a set $S$, 5 |
| $\alpha_{v_q} \oplus \beta_{v_q}$ | : the representation of a point in $T^*TQ$ in the splitting defined by a connection, 25 |
| $\operatorname{ann}(S)$ | : the annihilator of a subset $S$ of a vector space $V$, 5 |
| $\mathscr{A}$ | : the pair $(A, f)$ consisting of a symmetric $(0, r)$ tensor $A$ and a $C^\infty$ function $f$, 33 |
| $B_{\mathsf{Y}}$ | : a tensor associated with $\mathsf{Y}$, 41 |
| $C^\infty(M)$ | : the smooth functions on a manifold $M$, 5 |
| $\operatorname{Carc}(\Sigma)$ | : the set of controlled arcs for $\Sigma$, 29 |
| $\operatorname{Carc}(\Sigma, F)$ | : the set of $F$-acceptable controlled arcs for $\Sigma$, 29 |
| $\operatorname{Carc}(\Sigma, F, S_0, S_1)$ | : the $F$-acceptable controlled arcs which connect the submanifold $S_0$ to the submanifold $S_1$, 29 |

| | |
|---|---|
| $\mathrm{Carc}(\Sigma_{\mathrm{aff}})$ | : the set of controlled trajectories for $\Sigma_{\mathrm{aff}}$, 32 |
| $\mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}})$ | : the set of $F_{\mathscr{A}}$-acceptable controlled arcs for $\Sigma_{\mathrm{aff}}$, 34 |
| $\mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1)$ | : the $F_{\mathscr{A}}$-acceptable controlled trajectories which connect $q_0$ to $q_1$, 35 |
| $\mathrm{Carc}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1})$ | : the $F_{\mathscr{A}}$-acceptable controlled trajectories which connect $v_{q_0}$ to $v_{q_1}$, 35 |
| $\mathrm{Ctraj}(\Sigma)$ | : the set of controlled trajectories for $\Sigma$, 29 |
| $\mathrm{Ctraj}(\Sigma, F)$ | : the set of $F$-acceptable controlled trajectories for $\Sigma$, 29 |
| $\mathrm{Ctraj}(\Sigma_{\mathrm{aff}})$ | : the set of controlled trajectories for $\Sigma_{\mathrm{aff}}$, 32 |
| $\mathrm{Ctraj}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}})$ | : the set of $F_{\mathscr{A}}$-acceptable controlled trajectories for $\Sigma_{\mathrm{aff}}$, 34 |
| $\mathscr{D}(\Sigma)$ | : the set of valid inputs and states for $\Sigma$, 29 |
| $\mathscr{D}_Q(\Sigma_{\mathrm{aff}})$ | : the set of valid inputs and configurations for $\Sigma_{\mathrm{aff}}$, 32 |
| $\mathscr{D}_{TQ}(\Sigma_{\mathrm{aff}})$ | : the set of valid input and states for $\Sigma_{\mathrm{aff}}$, 33 |
| $F_{\mathscr{A}}$ | : the cost function for an affine connection control system associated with $\mathscr{A}$, 33 |
| $\mathscr{F}$ | : the collection $\{f_0, f_1, \dots, f_m\}$ of vector fields defining a control affine system, 28 |
| $\mathscr{F}(\Sigma_{\mathrm{aff}}, q_0, q_1)$ | : the force minimising, arbitrary interval, free velocity optimal control problem for an affine connection control system $\Sigma_{\mathrm{aff}}$, 40 |
| $\mathscr{F}(\Sigma_{\mathrm{aff}}, v_{q_0}, v_{q_1})$ | : the force minimising, arbitrary interval, fixed velocity optimal control problem for an affine connection control system $\Sigma_{\mathrm{aff}}$, 40 |
| $\gamma$ | : a typical controlled trajectory or controlled arc, 29, 32 |
| $\Gamma^\infty(E)$ | : the smooth sections of a vector bundle with total space $E$, 5 |
| $\Gamma^i_{jk}$ | : the Christoffel symbols for an affine connection, 12 |
| $H^{\Sigma,F}$ | : the Hamiltonian for the control affine system $\Sigma$ with cost function $F$, 30 |
| $H^{\Sigma,F}_{\min}$ | : the minimum Hamiltonian for the control affine system $\Sigma$ with cost function $F$, 30 |
| $H^{\Sigma,F}_u$ | : the Hamiltonian along a control $u$, 30 |
| $H^{\Sigma_{\mathrm{aff}},F_{\mathscr{A}}}$ | : the Hamiltonian for the affine connection control system $\Sigma_{\mathrm{aff}}$ with cost function $F_{\mathscr{A}}$, 36 |
| $H^{\Sigma_{\mathrm{aff}},F_{\mathscr{A}}}_{\min}$ | : the minimum Hamiltonian for the affine connection control system $\Sigma_{\mathrm{aff}}$ with cost function $F_{\mathscr{A}}$, 36 |
| $h_{\mathsf{Y}}$ | : the $(2,0)$ tensor field associated with $g$ and $\mathsf{Y}$, 40 |
| $\hat{A}$ | : constructs a one-form from a $(0,r)$ tensor field $A$ and $r-1$ vector fields, or constructs a vector field from a $(r,0)$ tensor $A$ and $r-1$ one-forms, 33, 34 |
| $HM$ | : an Ehresmann connection on a fibration $\pi\colon M \to B$, 16 |
| hor | : the horizontal projection associated with an Ehresmann connection, 16 |
| $\mathrm{hlft}(X)$ | : the horizontal lift of a vector field $X$, 16 |
| $I$ | : an interval in $\mathbb{R}$, possibly infinite or semi-infinite, 5 |
| $I_Q$ | : the canonical tangent bundle involution, 11 |

| | |
|---|---|
| $\mathrm{id}_S$ | : the identity map on a set $S$, 5 |
| $J^{\Sigma,F}$ | : the integrated cost function for a control affine system $\Sigma$ with cost function $F$, 29 |
| $J^{\Sigma_{\mathrm{aff}},F_{\mathscr{A}}}$ | : the integrated cost function for an affine connection control system $\Sigma_{\mathrm{aff}}$ with cost function $F_{\mathscr{A}}$, 34 |
| $J_M$ | : the canonical almost tangent structure on $TM$, 12 |
| $\mathscr{L}_X$ | : the Lie derivative with respect to a vector field $X$, 5 |
| LAC | : locally absolutely continuous, 6 |
| LAD | : locally absolutely differentiable, 6 |
| $\mathrm{LIC}^k$ | : locally integrable of class $C^k$ where $k \in \{0, 1, \dots, \infty\}$, 6 |
| $\nabla$ | : a general affine connection, 12 |
| $(\nabla X)^*$ | : the dual endomorphism of $\nabla X$, 13 |
| $\overset{g}{\nabla}$ | : the Levi-Civita affine connection associated with a Riemannian metric $g$, 13 |
| $\nabla\tau$ | : the covariant differential of an $(r, s)$ tensor field $\tau$, 12 |
| $\pi_{T^*M}$ | : the cotangent bundle projection for a manifold $M$, 5 |
| $\mathscr{P}(\Sigma, F, S_0, S_1)$ | : the arbitrary interval optimal control problem for a control affine system $\Sigma$, 29 |
| $\mathscr{P}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, q_0, q_1)$ | : the arbitrary interval, free velocity, optimal control problem for an affine connection control system $\Sigma_{\mathrm{aff}}$, 35 |
| $\mathscr{P}(\Sigma_{\mathrm{aff}}, F_{\mathscr{A}}, v_{q_0}, v_{q_1})$ | : the arbitrary interval, fixed velocity, optimal control problem for an affine connection control system $\Sigma_{\mathrm{aff}}$, 35 |
| $\mathscr{P}_{[a,b]}(\Sigma, F, S_0, S_1)$ | : the fixed interval optimal control problem for a control affine system $\Sigma$, 29 |
| $(q, p)$ | : natural coordinates for $T^*Q$, 5 |
| $(q, v)$ | : natural coordinates for $TQ$, 5 |
| $((q, v), (\alpha, \beta))$ | : natural coordinates for $T^*TQ$, 6 |
| $((q, v), (u, w))$ | : natural coordinates for $TTQ$, 6 |
| $R$ | : the curvature tensor for an affine connection, 14 |
| $R^*$ | : the "dual" of the curvature tensor $R$, 25 |
| $\overset{g}{R}$ | : the curvature tensor for the Levi-Civita connection associated with $g$, 14 |
| $\Sigma$ | : a general control affine system, 28 |
| $\Sigma_{\mathrm{aff}}$ | : a general affine connection control system, 31 |
| $T$ | : the torsion tensor of an affine connection, 14 |
| $T^*$ | : the "dual" of the torsion tensor $T$, 25 |
| $T^*M$ | : the cotangent bundle of a manifold $M$, 5 |
| $\mathrm{T}^r_s(E)$ | : the set of tensors of type $(r, s)$ on a vector bundle with total space $E$, 5 |
| $X^T$ | : the tangent lift of a vector field $X$, 7 |

| | |
|---|---|
| $X^{T^*}$ | : the cotangent lift of a vector field $X$, 8 |
| $T_x\phi$ | : the restriction of $T\phi$ to $T_xM$, 5 |
| $\pi_{TM}$ | : the tangent bundle projection for a manifold $M$, 5 |
| $TM$ | : the tangent bundle of a manifold $M$, 5 |
| $T\phi$ | : the derivative of a smooth map $\phi\colon M \to N$, 5 |
| $U$ | : the input range map for a control affine system or for an affine connection control system, 28, 31 |
| $u_{v_q} \oplus w_{v_q}$ | : the representation of a point in $TTQ$ in the splitting defined by a connection, 23 |
| ver | : the vertical projection associated with an Ehresmann connection, 16 |
| $\mathrm{vlft}(X)$ | : the vertical lift of a vector field $X$, 8 |
| $VM$ | : the vertical bundle of a fibration $\pi\colon M \to B$, 16 |
| $\omega$ | : the connection form associated with an Ehresmann connection, 16 |
| $\Omega$ | : the curvature form associated with an Ehresmann connection, 16 |
| $\mathscr{Y}$ | : the collection $\{Y_1, \dots, Y_m\}$ of control vector fields for an affine connection control system, 31 |
| $Z$ | : the geodesic spray associated with a general affine connection, 13 |

## References


Abraham, R., Marsden, J. E., and Ratiu, T. S. [1988] *Manifolds, Tensor Analysis, and Applications*, number 75 in Applied Mathematical Sciences, Springer-Verlag: New York/-Heidelberg/Berlin, ISBN: 978-0-387-96790-5.

Aliprantis, C. D. and Border, K. C. [2006] *Infinite-Dimensional Analysis, A Hitchhiker's Guide*, 3rd edition, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-3-540-32696-0.

Bloch, A. M. and Crouch, P. E. [1995] *Another view of nonholonomic mechanical control systems*, in *Proceedings of the 34th IEEE Conference on Decision and Control*, IEEE Conference on Decision and Control, (New Orleans, LA, Dec. 1995), Institute of Electrical and Electronics Engineers, pages 1066–1071, DOI: 10.1109/CDC.1995.480232.

Bullo, F. and Lynch, K. M. [2001] *Kinematic controllability and decoupled trajectory planning for underactuated mechanical systems*, Institute of Electrical and Electronics Engineers. Transactions on Robotics and Automation, **17**(4), pages 402–412, ISSN: 1042-296X, DOI: 10.1109/70.954753.

Chyba, M., Leonard, N. E., and Sontag, E. D. [2000] *Time-optimal control for underwater vehicles*, in *Proceedings of the 2000 IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control, (Princeton, NJ, Mar. 2000), edited by N. E. Leonard and R. Ortega, International Federation of Automatic Control, pages 122–127.

Coombs, A. T. [2000] "Time-Optimal Control for Two Simple Mechanical Systems with Three Degrees of Freedom and Two Inputs", Masters thesis, Kingston, ON, Canada: Queen's University, Kingston.

Crampin, M. [1983] *Tangent bundle geometry for Lagrangian dynamics*, Journal of Physics. A. Mathematical and Theoretical, **16**(16), pages 3755–3772, ISSN: 1751-8113, DOI: 10.1088/0305-4470/16/16/014.

Crouch, P. E. and Silva Leite, F. [1991] *Geometry and the dynamic interpolation problem*, in *Proceedings of the 1991 American Control Conference*, American Control Conference, (Boston, MA, June 1991), Institute of Electrical and Electronics Engineers, pages 1131–1136.

— [1995] *The dynamic interpolation problem: On Riemannian manifolds, Lie groups, and symmetric spaces*, Journal of Dynamical and Control Systems, **1**(2), pages 177–202, ISSN: 1079-2724, DOI: 10.1007/BF02254638.

Crouch, P. E., Silva Leite, F., and Camarinha, M. [2000] *Hamiltonian structure of generalized cubic polynomials*, in *Proceedings of the 2000 IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control, (Princeton, NJ, Mar. 2000), edited by N. E. Leonard and R. Ortega, International Federation of Automatic Control, pages 13–18.

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].

— [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].

Kolář, I., Michor, P. W., and Slovák, J. [1993] *Natural Operations in Differential Geometry*, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-3-540-56235-1.

Koon, W.-S. and Marsden, J. E. [1997] *Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction*, SIAM Journal on Control and Optimization, **35**(3), pages 901–929, ISSN: 0363-0129, DOI: 10.1137/S0363012995290367.

Lewis, A. D. [2000] *Simple mechanical control systems with constraints*, Institute of Electrical and Electronics Engineers. Transactions on Automatic Control, **45**(8), pages 1420–1436, ISSN: 0018-9286, DOI: 10.1109/9.871752.

Lewis, A. D. and Murray, R. M. [1997] *Configuration controllability of simple mechanical control systems*, SIAM Journal on Control and Optimization, **35**(3), pages 766–790, ISSN: 0363-0129, DOI: 10.1137/S0363012995287155.

Liu, W. S. and Sussmann, H. J. [1994] *Abnormal sub-Riemannian minimizers*, in *Differential Equations, Dynamical Systems, and Control Science*, Festschrift in honor of Lawrence Markus, (1994), edited by K. D. Elworthy, W. N. Everitt, and E. B. Lee, 152 Lecture Notes in Pure and Applied Mathematics, pages 705–716, Dekker Marcel Dekker: New York, NY, ISBN: 978-0-8247-8904-6.

Montgomery, R. [1994] *Abnormal minimizers*, SIAM Journal on Control and Optimization, **32**(6), pages 1605–1620, ISSN: 0363-0129, DOI: 10.1137/S0363012993244945.

Noakes, L., Heinzinger, G., and Paden, B. [1989] *Cubic splines on curved spaces*, IMA Journal of Mathematical Control and Information, **6**(4), pages 465–473, ISSN: 0265-0754, DOI: 10.1093/imamci/6.4.465.

Ostrowski, J. P., Desai, J. P., and Kumar, V. [1999] *Optimal gait selection for nonholonomic locomotion systems*, The International Journal of Robotics Research, **19**(2), pages 225–237, ISSN: 0278-3649, DOI: 10.1177/02783640022066833.

Silva Leite, F., Camarinha, M., and Crouch, P. E. [2000] *Elastic curves as solutions of Riemannian and sub-Riemannian control problems*, Mathematics of Control, Signals, and Systems, **13**(2), pages 140–155, ISSN: 0932-4194, DOI: 10.1007/PL00009863.

Sontag, E. D. [1998] *Mathematical Control Theory, Deterministic Finite Dimensional Systems*, 2nd edition, number 6 in Texts in Applied Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-98489-6.

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.

Synge, J. L. [1928] *Geodesics in nonholonomic geometry*, Mathematische Annalen, **99**(1), pages 738–751, ISSN: 0025-5831, DOI: 10.1007/BF01459122.

Tulczyjew, W. M. [1977] *The Legendre transformation*, Annales de l'Institut Henri Poincaré. Physique Théorique, **27**(1), pages 101–114, ISSN: 0246-0211, URL: http://www.numdam.org/item?id=AIHPA_1977__27_1_101_0 (visited on 07/14/2014).

Yano, K. and Ishihara, S. [1973] *Tangent and Cotangent Bundles*, number 16 in Pure and Applied Mathematics, Dekker Marcel Dekker: New York, NY, ISBN: 978-0-8247-6071-7.