Energy-preserving affine connections

Andrew D. Lewis†
28/07/1997

Abstract
A Riemannian affine connection on a Riemannian manifold has the property that is preserves the “kinetic energy” associated with the metric. However, there are other affine connections which have this property, and here we characterise them. A class of such energy-preserving affine connections, related to mechanical systems with constraints, is provided.

Keywords. mechanics, conservation of energy, affine connections, distributions

AMS Subject Classifications. 53B20, 55C22, 70H03

1. Introduction

On a Riemannian manifold \((M,g)\), geodesics of the Levi-Civita affine connection, or more generally, a Riemannian affine connection (i.e., not necessarily torsion-free) leave invariant the “kinetic energy” of the Riemannian metric. Precisely, if \(Z\) is the geodesic spray of a Riemannian connection \(\nabla\), and if we define a function \(K(v) = \frac{1}{2}g(v,v)\) on \(TM\), then \(\mathcal{L}_Z K = 0\) where \(\mathcal{L}_Z\) is the Lie derivative with respect to \(Z\). It is interesting to consider to what extent this characterises energy-preserving affine connections. Here we provide analytical conditions for a general affine connection to leave invariant the kinetic energy.

Motivated by a construction of Synge [1928], given a distribution \(D\) on \(M\), we construct a natural energy-preserving affine connection on \(M\) which restricts to \(D\). This construction has been explored in detail by the author in [Lewis 1998]. Applications to control theory are provided in [Lewis 2000].

In this paper we will follow the differential geometric notation and conventions of Abraham, Marsden, and Ratiu [1988], and we refer to [Kobayashi and Nomizu 1963] for background on affine connections. We denote by \(\mathcal{C}^\infty(M)\) the set of \(C^\infty\) functions on a manifold \(M\), by \(\mathcal{F}(M)\) the \(C^\infty\) vector fields, and by \(\mathcal{F}^*(M)\) the \(C^\infty\) one-forms. If \((M,g)\) is a Riemannian manifold, we denote by \(\nabla\) the Levi-Civita affine connection. We recall that in coordinates the Christoffel symbols of \(\nabla\) are

\[
\Gamma^i_{jk} = \frac{1}{2} \partial_{\alpha} \left( g^{il} \partial_j g_{\alpha l} + \partial_j g_{\alpha l} + \partial_k g_{\alpha l} \right).
\]

(1.1)

If \(S\) is an \((r,s)\) tensor field and \(\nabla\) is an arbitrary affine connection, we denote by \(\nabla S\) the \((r,s+1)\) tensor field defined by

\[
\nabla S(a^{1}, \ldots, a^{r}, X_0, X_1, \ldots, X_s) = \nabla_{X_0} S(a^{1}, \ldots, a^{r}, X_1, \ldots, X_s).
\]

2. Characterising energy-preserving affine connections

Let \((M,g)\) be a finite-dimensional Riemannian manifold. If \(c : [a, b] \to M\) is a curve on \(M\), define a function \(E_c\) along \(c\) by \(E_c(t) = \frac{1}{2}g(\dot{c}(t), \dot{c}(t))\). An affine connection \(\nabla\) on \(M\) is energy-preserving if for every geodesic \(c\) of \(\nabla\) we have \(\frac{d}{dt} E_c = 0\). One may readily verify that this is equivalent to saying, as we did in the introduction, that \(\mathcal{L}_Z K = 0\) where \(Z\) is the geodesic spray of \(\nabla\) and \(K(v) = \frac{1}{2}g(v,v)\). The definition we give here is better suited to the computations we will here perform.

To motivate our first result, let us perform a coordinate computation. The computation we perform comes from one made in [Murray, Li, and Sastry 1994, Lemma 4.2.1] to derive stability results for certain control laws in robotics. We let \(M = \mathbb{R}^n\) and let \(g\) be a Riemannian metric on \(\mathbb{R}^n\). In standard coordinates \((x^1, \ldots, x^n)\) for \(\mathbb{R}^n\) we have \(E_c = \frac{1}{2}g_{ij} \dot{x}^i \dot{x}^j\) for a curve \(c : t \mapsto (x^1(t), \ldots, x^n(t))\). If \(c\) is a geodesic of \(\nabla\) then \(\dot{x}^i = -\Gamma^i_{jk} \dot{x}^j \dot{x}^k\). Now, for a geodesic \(c\) of \(\nabla\) we compute

\[
\frac{d}{dt} E_c = g_{ij} \ddot{x}^i \dot{x}^j + \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \dot{x}^j \dot{x}^k = \frac{1}{2} \left( \frac{\partial g_{ij}}{\partial x^k} - 2g_{ik} \Gamma^l_{jk} \right) \dot{x}^j \dot{x}^k.
\]

(2.1)

Using (1.1) we readily compute

\[
\frac{\partial g_{ij}}{\partial x^k} - 2g_{ik} \Gamma^l_{jk} = \frac{\partial g_{lj}}{\partial x^k} - \frac{\partial g_{lk}}{\partial x^j}.
\]

(2.2)

Skew-symmetry of this expression in the indices \(i\) and \(j\) immediately implies that \(\frac{d}{dt} E_c = 0\). This raises the question which we address in this paper; namely the question of exactly when does an affine connection preserve energy.

It is helpful to have the following result from multilinear algebra. We denote by \(S_\sigma\) the permutation group on \(k\) symbols. If \(A\) is a \((0,k)\) tensor on a vector space \(V\) define the symmetric part of \(A\) by

\[
\text{Sym}(A)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_\sigma} A(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).
\]

Of course, the same construction applies to \((0,k)\) tensor fields on manifolds.

2.1 Lemma: Let \(V\) be a vector space over a field of characteristic zero and let \(A\) be a \((0,k)\) tensor on \(V\). The following are equivalent:

(i) \(A(v_1, \ldots, v_k) = 0\) for every \(v_i \in V\);

(ii) \(\text{Sym}(A) = 0\).

Proof: (i) \(\implies\) (ii) Let \(A\) be a \((0,k)\) tensor with the property that \(A(v_1, \ldots, v_k) = 0\) for all \(v_i \in V\). Using the latter property of \(A\), for \(v_1, \ldots, v_k \in V\) we have

\[
A \left( \sum_{i_1=1}^{k} v_{i_1}, \ldots, \sum_{i_k=1}^{k} v_{i_k} \right) = \sum_{\sigma \in S_k} A(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) + \sum_{l=1}^{k-1} \sum_{\{i_1,\ldots,i_l\} \subset\{1,\ldots,k\}} A(v_{i_1}, \ldots, v_{i_l}, v_{j_1}, \ldots, v_{j_l}).
\]
where the inner sum in the final term is over all distinct subsets \( j_1, \ldots, j_l \subset \{ 1, \ldots, k \} \) with \( j_1 < \ldots < j_l < \ldots < j_l \). This then gives
\[
\sum_{\sigma \in S_k} A(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = 0
\]
for \( v_1, \ldots, v_k \in V \) and so \( \text{Sym}(A) = 0 \).

(ii) \( \iff \) (i) For \( v \in V \) we have
\[
\sum_{i=1}^k A(v, \ldots, v) = kA(v, \ldots, v) = 0
\]
so \( A(v, \ldots, v) = 0 \) since \( V \) is over a field of characteristic zero.

Now we can easily give the following characterisation of energy-preserving affine connections.

2.2 Proposition: Let \((M, g)\) be a Riemannian manifold and let \( \nabla \) be an arbitrary affine connection on \( M \). The following are equivalent:

(i) \( \nabla \) is energy-preserving;

(ii) \( \text{Sym}(\nabla g) = 0 \);

(iii) for every chart \((U, \phi)\) for \( M \) with coordinates \((x^1, \ldots, x^n)\) we have
\[
\sum_{\sigma \in S_k} \frac{\partial g_{\sigma(i)\sigma(j)}}{\partial x^{\sigma(k)}} - 2g_{\sigma(i)\sigma(j)}\Gamma_{\sigma(k)\sigma(i)\sigma(j)} = 0
\]
for \( i_1, j_1, k_1 = 1, \ldots, n \). Here \( \Gamma_{\sigma(k)\sigma(i)\sigma(j)} \) are the Christoffel symbols of \( \nabla \) in the chart \((U, \phi)\).

Proof: (i) \( \iff \) (ii) The affine connection \( \nabla \) is energy-preserving if and only if for every geodesic \( c: [a, b] \to M \) of \( \nabla \) we have
\[
\frac{d}{dt} E_i(t) = \frac{1}{2} \nabla_{\dot{c}(t)}(g(\dot{c}(t), \dot{c}(t))) = \frac{1}{2} \nabla_{\dot{c}(t)}(\dot{c}(t), \dot{c}(t)) = 0.
\]
Thus \( \nabla \) is energy-preserving if and only if \( \nabla g(v, v, v) = 0 \) for every \( v \in T_c M \) and \( c \in M \). The result now follows from Lemma 2.1.

(i) \( \iff \) (iii) Proceeding exactly as we did in the computation of (2.1) we see that \( \nabla \) is energy-preserving if and only if
\[
\left( \frac{\partial g_{ij}}{\partial x^k} - 2g_{ij}\Gamma^k_{ij} \right)(x)v^i v^j = 0
\]
for every \( x \in M \), every chart \((U, \phi)\) around \( x \) with coordinates \((x^1, \ldots, x^n)\), and every \( v = v^i \partial_i \) in \( T_x M \). The result now follows from Lemma 2.1.

This shows, in particular, that if
\[
\frac{\partial g_{ij}}{\partial x^k} - 2g_{ij}\Gamma^k_{ij} = -\frac{\partial g_{ij}}{\partial x^k} + 2g_{ij}\Gamma^k_{ij}
\]
then \( \nabla \) is energy-preserving. In particular, by (2.2) this means that the Levi-Civita affine connection is energy preserving. It is also clear that any Riemannian affine connection is energy preserving. Indeed, a Riemannian affine connection \( \nabla \) is characterised by \( \nabla g = 0 \), so (ii) shows that \( \nabla \) is energy-preserving.

3. A class of energy-preserving affine connections

In this section we provide a collection of non-Riemannian energy-preserving affine connections. These arise in a natural way if one is given a distribution \( D \) on \( M \). First we review some of the ideas of Lewis [1998].

3.1 Affine connections and distributions. Let \( \nabla \) be an affine connection on \( M \) and let \( D \) be a distribution on \( M \). We denote by \( \mathcal{D} \) the set of vector fields on \( M \) taking their values in \( D \). We say that \( \nabla \) restricts to \( D \) if \( \nabla_X Y \in \mathcal{D} \) for every \( Y \in \mathcal{D} \). Thus, if \( \nabla \) restricts to \( D \) then we may regard \( \nabla \) as a connection in the vector bundle \( D \). It is important that one not require an affine connection which restricts to a distribution to be torsion-free; one readily sees that this implies that \( D \) is integrable.

A notion weaker than restricting to a distribution is that of geodesic invariance. We shall say \( D \) is \textit{geodesically invariant} if for every geodesic \( c: [a, b] \to M \) of \( \nabla \), \( \dot{c}(t) \in D_{c(t)} \) implies that \( \dot{c}(t) \in D_{c(t)} \) for \( t \in [a, b] \). The following result characterises geodesically invariant distributions.

3.1 Theorem: ([Lewis [1998]]) \( D \) is geodesically invariant if and only if \( \nabla_X Y + \nabla_Y X \in \mathcal{D} \) for every \( X, Y \in \mathcal{D} \).

3.2 Restricting the Levi-Civita connection to a distribution. Let \((M, g)\) be a Riemannian manifold and let \( D \) be a distribution on \( M \). Let \( D^x \) denote the orthogonal complement of \( D \) and denote by \( P: TM \to TM \) and \( P^*: TM \to TM \) the orthogonal projections onto \( D \) and \( D^x \), respectively. If \( v: TM \to T^*M \) and \( j: TM \to TM \) are the musical isomorphisms, we define \( \hat{\nabla} = D^x = \{ D^x \} \) and \( \hat{\nabla}^x = D = \{ D \} \), the codistributions of annihilators of \( D^x \) and \( D \), respectively. We write the projections onto \( D \) and \( D^x \) as \( \hat{P}: T^*M \to T^*M \) and \( \hat{P}^*: T^*M \to T^*M \), respectively.

If \( \hat{\nabla} \) denotes the Levi-Civita connection, we wish to define a new affine connection \( \hat{\nabla} \) on \( M \) which has the properties

1. \( \hat{\nabla} \) restricts to \( D \), and
2. \( \nabla_X Y - \hat{\nabla}_X Y \in \mathcal{D}^x \) for \( Y \in \mathcal{D}^x \).

Stated otherwise, \( \nabla_X Y = \hat{P}(\hat{\nabla}_X Y) \) for \( Y \in \mathcal{D}^x \). Vershik [1984] defines \( \hat{P}(\hat{\nabla}_X Y) \) as a vector bundle connection, and makes the point that there is no natural way to extend this to an affine connection on \( M \). Here we describe what such an extension should look like. The following result is given in [Lewis 1998].

3.2 Proposition: Let \( D \) be a distribution on a Riemannian manifold \((M, g)\). Let \( \hat{\nabla} \) be the Levi-Civita connection and suppose that another affine connection \( \nabla \) has the properties

(i) \( \nabla_X Y \in \mathcal{D}^x \) for every \( Y \in \mathcal{D}^x \), and

(ii) \( \nabla_X Y - \hat{\nabla}_X Y \in \mathcal{D}^x \) for every \( Y \in \mathcal{D}^x \).

Then \( \nabla_X Y = \hat{P}(\hat{\nabla}_X Y) + S(X, Y) \) for some \((1,2)\) tensor field \( S \) such that \( P(S(X, Y)) = 0 \) for \( Y \in \mathcal{D}^x \). Conversely, if \( \nabla \) is of this form, then it satisfies (i) and (ii).
Proof: We may write any affine connection on $M$ as

$$\nabla_X Y = \tilde{\nabla}_X Y + B(X, Y)$$

for some (1, 2) tensor field $B$. In particular, an affine connection satisfying (i) and (ii) must be of this form. For $Y \in \mathcal{D}$ and any vector field $X$ we have

$$P'(Y) = 0 \implies (\tilde{\nabla}_X P')(Y) + P'(\tilde{\nabla}_X Y) = 0$$
$$\implies P'(\tilde{\nabla}_X Y) = -P'(\tilde{\nabla}_X P')(Y). \quad (3.1)$$

We also have

$$\nabla_X Y = \tilde{\nabla}_X Y + B(X, Y).$$

Using (i), (ii), and (3.1) we have

$$P'(\tilde{\nabla}_X Y) + B(X, Y) = 0 \implies B(X, Y) = (\tilde{\nabla}_X P')(Y).$$

Thus $\nabla_X Y = \tilde{\nabla}_X Y + (\tilde{\nabla}_X P')(Y) + S(X, Y)$ for some $S$ such that $P'(S(X, Y)) = 0$ for $Y \in \mathcal{D}$.

Now we show that $\nabla$ satisfies (i) and (ii) if it is of the given form. Let $X$ and $Y$ be vector fields on $M$. Then

$$P'(\tilde{\nabla}_X Y) = P'(\tilde{\nabla}_X Y) + P'(\tilde{\nabla}_X P')(Y). \quad (3.2)$$

If $Y \in \mathcal{D}$ then

$$P'(Y) = 0 \implies (\tilde{\nabla}_X P')(Y) + P'(\tilde{\nabla}_X Y) = 0$$
$$\implies P'(\tilde{\nabla}_X Y) + P'(\tilde{\nabla}_X P')(Y) = 0 \quad (3.3)$$

since $P' P' = P'$. Substituting (3.4) into (3.2) we see that $P'(\tilde{\nabla}_X Y) = 0$ for $X \in \mathcal{D}(M)$ and $Y \in \mathcal{D}$. Therefore, $\nabla_X Y \in \mathcal{D}$. Now let $Y \in \mathcal{D}$. From (3.3) we have

$$(\tilde{\nabla}_X P')(Y) + P'(\tilde{\nabla}_X Y) = 0$$
$$\implies P'(\tilde{\nabla}_X P')(Y) = 0 \quad (3.4)$$

since $P' P' = P'$. Thus $(\tilde{\nabla}_X P')(Y) \in \mathcal{D}$ for $Y \in \mathcal{D}$.

We will also find it useful to have the following formula:

$$(\tilde{\nabla}_X P')(Y) \in \mathcal{D}, \quad Y \in \mathcal{D}.$$  \quad (3.5)$$

This may be verified as follows. Let $Y \in \mathcal{D}$. Then

$$P'(Y) = Y$$
$$\implies (\tilde{\nabla}_X P')'(Y) + P'(\tilde{\nabla}_X Y) = \tilde{\nabla}_X Y$$
$$\implies P'(\tilde{\nabla}_X P')(Y) + P'(\tilde{\nabla}_X Y) = P'(\tilde{\nabla}_X Y)$$
$$\implies P'(\tilde{\nabla}_X P')(Y) = 0.$$

Let us see how this relates to nonholonomic mechanics. We consider a mechanical system on $M$ with Lagrangian $L_v = \frac{1}{2}g(v, v)$. The distribution $D$ defines a nonholonomic constraint for the mechanical system. Thus we require the velocities of the system to lie in $D$. The Lagrange-d’Alembert principle for constrained systems states that solutions $c: [a, b] \to M$ of the constrained problem satisfy

$$\tilde{\nabla}_{\dot{c}}(\dot{c})(t) = \lambda(t)$$
$$P'(\dot{c})(t) = 0 \quad (3.6a)$$

where $\lambda$ is a section of $D^\perp$ along $c$ (a Lagrange multiplier, if you like). We now show how solutions to (3.6) are related to the affine connections described by Proposition 3.2. This is an intrinsic version of the construction of Synge [1928].

3.3 Proposition: A curve $c: [a, b] \to M$ is a solution to (3.6) if and only if $\dot{c}(a) \in D_{(a)}$ and $c$ is a geodesic of any of the affine connections characterised by Proposition 3.2.

Proof: Proposition 3.2(i) and Theorem 3.1 imply that $D$ is geodesically invariant. Thus, if $\nabla$ is one of the affine connections characterised by Proposition 3.2, every geodesic of $\nabla$ whose initial velocity lies in $D$ will satisfy (3.6b).

Now suppose that $c$ is a solution of (3.6) and let $\nabla$ be an affine connection as in Proposition 3.2. We may differentiate (3.6b) to obtain

$$(\tilde{\nabla}_{\dot{c}}(\dot{c}))(t) + P'(\tilde{\nabla}_{\dot{c}}(\dot{c}))(t) = 0.$$

From (3.6a) we also have $P'(\tilde{\nabla}_{\dot{c}}(\dot{c}))(t) = \lambda(t)$ since $\lambda$ is a section of $D^\perp$. We then see that

$$\lambda(t) = -(\tilde{\nabla}_{\dot{c}}(\dot{c}))(t)$$
$$\implies \dot{c} = -(\tilde{\nabla}_{\dot{c}}(\dot{c}))(t)$$

so that $c$ satisfies

$$\tilde{\nabla}_{\dot{c}}(\dot{c})(t) + (\tilde{\nabla}_{\dot{c}}(\dot{c}))(t) = 0.$$

If we define a (1, 2) tensor field $S$ by $S(X, Y) = \nabla_X Y - \tilde{\nabla}_X Y - (\nabla_X P')(Y)$ then we have

$$\tilde{\nabla}_{\dot{c}}(\dot{c})(t) + (\tilde{\nabla}_{\dot{c}}(\dot{c}))(t) + S(\dot{c}(t), \dot{c}(t)) = 0$$

by the properties of $\nabla$. Thus $c$ is a geodesic of $\nabla$.

Now suppose $c: [a, b] \to M$ is a geodesic of an affine connection $\nabla$ defined by Proposition 3.2 and that $\dot{c}(a) \in D_{(a)}$. Then $c$ satisfies (3.6b) by Proposition 3.2(i). Also, $c$ satisfies (3.6a) with $\lambda$ defined by (3.7). This completes the proof.
3.3. Energy-preserving affine connections which restrict to a distribution. In this section we combine all that we have done in previous sections and construct an affine connection of the type described by Proposition 3.2 and which preserves energy. The following lemma will be useful.

3.4 Lemma: Let $(M, g)$ be a Riemannian manifold with $\hat{\nabla}$ the Levi-Civita affine connection and $\nabla$ another affine connection defined by

$$\nabla_x Y = \hat{\nabla}_x Y + A(x, Y)$$

for a $(1, 2)$ tensor field $A$. For vector fields $X, Y, Z$ on $M$ we have

$$(\nabla_x g)(Y, Z) = -g(A(x, Y), Z) - g(Y, A(x, Z)).$$

Proof: We compute

$$L_x(g(Y, Z)) = (\hat{\nabla}_x g)(Y, Z) + g(\hat{\nabla}_x Y, Z) + g(Y, \hat{\nabla}_x Z).$$

Also

$$L_x(g(Y, Z)) = (\nabla_x g)(Y, Z) + g(\nabla_x Y, Z) + g(Y, \nabla_x Z)$$

$$= (\nabla_x g)(Y, Z) + g(\hat{\nabla}_x Y, Z) + g(Y, \hat{\nabla}_x Z) +$$

$$g(A(x, Y), Z) + g(Y, A(x, Z)).$$

Subtracting (3.8) from (3.9) and using the fact that $\hat{\nabla}x g = g$ we obtain the desired result.

To characterise the affine connections we are interested in, it will be convenient to introduce some notation. If $A$ is an $(r, s)$ tensor field on a Riemannian manifold $(M, g)$ then we define an $(0, r + s)$ tensor field $\tilde{A}$ by

$$A(x_1, \ldots, x_r, x_{r+1}, \ldots, x_{r+s}) = A(X_1, \ldots, X_r, X_{r+1}, \ldots, X_{r+s}).$$

With this we state the following result.

3.5 Proposition: An affine connection $\nabla: (X, Y) \mapsto \hat{\nabla}_x Y + (\hat{\nabla}_x P') Y + S(x, Y)$ of the form specified by Proposition 3.2 is energy-preserving if and only if

$$\text{Sym}(\hat{\nabla}) + \text{Sym}(\hat{\nabla} P) = 0.$$

Proof: Apply Proposition 2.2(ii) and Lemma 3.4.

Of course this does not guarantee the existence of energy-preserving affine connections of the form specified by Proposition 3.2. The following result, however, gives an example of just such an affine connection.

3.6 Proposition: Let $(M, g)$ be a Riemannian manifold with $\cal D$ a distribution on $M$. There exists an affine connection $\nabla$ on $M$ with the following properties:

(i) $\nabla$ satisfies (i) and (ii) of Proposition 3.2;

(ii) $\nabla$ is energy-preserving.

Furthermore, the affine connection $\hat{\nabla}$ defined by

$$\hat{\nabla}_x Y = \hat{\nabla}_x Y + \hat{\nabla}_x P' (Y) + S(x, Y).$$

for $X, Y \in \mathcal{F}(M)$ and $a \in \mathcal{F}^*(M)$, satisfies these properties.

Proof: We note that

$$\hat{\nabla}_x P' (Y) = \hat{\nabla}_x P'(Y) + \hat{\nabla}_x P'(P(Y)) +$$

$$\hat{\nabla}_x P'(P(Y)) + \hat{\nabla}_x P'(P(Y)) =$$

$$\hat{\nabla}_x P'(P(Y)) + \hat{\nabla}_x P'(P(Y))$$

by Proposition 3.2(ii) and (3.5). Thus $\hat{\nabla}$ can be defined by

$$\hat{\nabla}_x P'(Y) = \hat{\nabla}_x P'(Y) - \hat{\nabla}_x P'(P(Y)) - \hat{\nabla}_x P'(P(Y)).$$

By Proposition 3.2, $\hat{\nabla}$ satisfies (i). Next we show that $\hat{\nabla}$ is energy preserving. Let $c$ be a geodesic of $\hat{\nabla}$ and compute

$$\frac{d}{dt} E_c(t) = \frac{1}{2} \nabla \alpha (\nabla \alpha (g)) \dot{c}(t), \dot{c}(t) = 0$$

using Lemma 3.4 and the fact that $c$ is a geodesic of $\nabla$. This establishes the existence of an affine connection with the specified properties.

3.7 Corollary: For every solution $c$ of (3.8), $\frac{d}{dt} E_c = 0$. That is, energy is conserved for constrained systems.

The affine connection $\hat{\nabla}$ given in Proposition 3.6 is distinguished in another way than its being energy-preserving. Recall that if $\nabla$ is an arbitrary affine connection on $M$, a diffeomorphism $\phi: M \to M$ is an affine transformation for $\nabla$ if $\phi^*(\nabla_x Y) = \nabla_{\phi x}^X \phi Y$ for every $X, Y \in \mathcal{F}(M)$. A diffeomorphism $\phi$ is compatible with a distribution $D$ if $T_0(\phi X) = D(\phi X)$ for every $x \in M$.

3.8 Proposition: If a diffeomorphism $\phi: M \to M$ is an affine transformation for $\nabla$ and is compatible with $D$, then $\phi$ is an affine transformation for $\hat{\nabla}$.

Proof: We first claim that $\phi^*(\nabla_x Y) = \hat{\nabla}_x Y$ for every $X \in \mathcal{F}(M)$. Indeed, if we write $X = X_1 + X_2$ where $X_1 \in \mathcal{D}$ and $X_2 \in \mathcal{D}^\perp$ we have $P(X) = X_1$. Also, $\phi^* X_1 \in \mathcal{D}$ since $\phi$ is compatible with $D$. For any $Y_1 \in \mathcal{D}$ and $Y_2 \in \mathcal{D}^\perp$ we compute

$$0 = \phi^* g(Y_1, Y_2) = \phi^* g(\phi^* Y_1, \phi^* Y_2) = g(\phi^* Y_1, \phi^* Y_2)$$
which shows, in particular, that \( \phi^* X_2 \in \mathcal{D}^\perp \). Therefore, \( P(\phi^* X) = \phi^* X_1 \). This demonstrates our claim that \( \phi^*(P(X)) = P(\phi^* X) \) for every \( X \in \mathcal{T}(M) \). In like fashion we have \( \phi^*(P(X)) = P(\phi^* X) \) for every \( X \in \mathcal{T}(M) \). From this it also follows that \( \phi^* P = P' \).

We now claim that \( \phi^* \hat{\nabla} P' = \hat{\nabla} P' \). For \( X, Y \in \mathcal{T}(M) \) and \( \alpha \in \mathcal{F}^*(M) \) we have

\[
\phi^*(\hat{\nabla} P'(\phi, \alpha, \phi X, \phi Y)) = (\phi^* \hat{\nabla} P') (\alpha, X, Y).
\]

We also have

\[
\phi^*(\hat{\nabla}_{\phi X} P'(\phi, \alpha, \phi X, \phi Y)) = \phi^*(\hat{\nabla}_{\phi X} P'(\phi, \alpha, \phi X, \phi Y)) - \phi^*(\hat{\nabla}_{\phi X} P'(\phi, \alpha, \phi Y)) = \hat{\nabla}_X (P'(Y)) - P' (\hat{\nabla}_X Y) = (\hat{\nabla} X P')(Y)
\]

using the fact that \( \phi \) is an affine transformation for \( \hat{\nabla} \) and that \( \phi^* P' = P' \). This then gives

\[
\phi^*(\hat{\nabla} P'(\phi, \alpha, \phi X, \phi Y)) = \alpha (\phi^*(\hat{\nabla}_{\phi X} P')(\phi, \alpha, \phi Y)) = \alpha (\hat{\nabla}_X P')(Y) = \hat{\nabla} P'(\alpha, X, Y).
\]

This shows that \( (\phi^* \hat{\nabla} P')(\alpha, X, Y) = \hat{\nabla} P'(\alpha, X, Y) \) for every \( X, Y \in \mathcal{T}(M) \) and \( \alpha \in \mathcal{F}^*(M) \). In other words, \( \phi^* \hat{\nabla} P' = \hat{\nabla} P' \).

The result now follows easily from the above computations and the definition of \( \hat{\nabla} \).

Lewis [1998] discusses “\( D \)-affine transformations” by restricting interest only to \( D \). Here we have shown that these notions can be extended to all of \( TM \).

Acknowledgements

The author wishes to thank Richard Murray for a few years ago asking the question, “What is the meaning of \( M - 2C' \)” (the quantity \( M - 2C \) being roughly the non-intrinsic object described in coordinates by (2.2)).

References


— [2000] Simple mechanical control systems with constraints, Institute of Electrical and Electronics Engineers, Transactions on Automatic Control, 45(8), 1420–1436.


A. D. Lewis
