Reports on Mathematical Physics
42(1/2), 135–164, 1998
DOI: 10.1016/S0034-4877(98)80008-6
Appears as *Affine connections and distributions with applications to nonholonomic mechanics*

Affine connections and distributions

Andrew D. Lewis*

1997/09/18
Last updated: 2002/03/06

Abstract

We investigate various aspects of the interplay of an affine connection with a distribution. When the affine connection restricts to the distribution, we discuss torsion, curvature, and holonomy of the affine connection. We also investigate transformations which respect both the affine connection and the distribution. A stronger notion than that of restricting to a distribution is that of geodesic invariance. This is a natural generalisation to a distribution of the idea of a totally geodesic submanifold. We provide a product for vector fields which allows one to test for geodesic invariance in the same way one uses the Lie bracket to test for integrability. If the affine connection does not restrict to the distribution, we are able to define its restriction and in the process generalise the notion of the second fundamental form for submanifolds. We characterise some transformations of these restricted connections and derive conservation laws in the case when the original connection is the Levi-Civita connection associated with a Riemannian metric.

Keywords. affine connections, distributions, geodesic invariance, symmetric product.

AMS Subject Classifications (2020). 53B05, 53B25, 53C22.

1. Introduction

In this paper we study relationships between distributions and affine connections. In large part, what we study here generalises what one normally does for submanifolds. If an affine connection restricts to a distribution, then it is interesting to study the affine connection thought of as a connection only in the distribution. We demonstrate that this is feasible by showing that if the affine connection restricts to the distribution, then so does itscurvature and holonomy. We also investigate transformations of affine connections which restrict to a distribution. One is then interested in transformations which respect both the affine connection *and* the distribution. A notion which is weaker than restriction to a distribution is what we call ***geodesic invariance***. A distribution D is geodesically invariant if $D \subset TM$ is invariant under the geodesic flow. We provide an infinitesimal test for geodesic invariance using the ***symmetric product***. This product was first seen in the Levi-Civita context in the work of Crouch [1981] on gradient control systems. It also arises in the work of Lewis and Murray [1997] on a class of mechanical control systems.

*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 performed while Research Fellow at the California Institute of Technology.
Research in part supported by NSF Grant CMS–9502224.

Here we provide a geometric interpretation of the symmetric product. Another notion from submanifold geometry which we generalise to distributions is the second fundamental form. By restricting an affine connection to a distribution in a natural way, we obtain an explicit formula for the second fundamental form for distributions and we show that it does indeed generalise the existing notion for submanifolds. Using our discussion of transformations for affine connections which restrict to a distribution, we provide some transformations for the restricted affine connections constructed. In the Levi-Civita case, certain of these transformations lead to conservation laws which may be thought of as generalisations of Noether's theorem.

We review some concepts from the theory of affine and linear connections in Section 2 in order to fix our notation. In Section 3 we study properties of affine connections which restrict to distributions. The torsion, curvature, and holonomy of such affine connections are discussed in Section 3.1 and their transformations are discussed in Section 3.2. We discuss geodesic invariance and the symmetric product in Section 4. Conditions are provided under which an affine connection which possesses D as a geodesically invariant distribution actually restricts to D. In Section 5 we look at the situation when we have an *arbitrary* affine connection ∇ and a distribution D. We discuss the second fundamental form in Section 5.2 and transformations in Section 5.3. The conservation laws in the Levi-Civita case are presented in Section 5.4. A simple example illustrates some of our constructions in Section 5.6.

Notation and conventions

In the rest of the paper we use the following notation.

$C^\infty(M)$	: the set of smooth functions on M
$\boldsymbol{D}\phi$	: the derivative of $\phi\colon U \Subset E \to F$
$\mathscr{D}$	: the set of smooth sections of a distribution D
E, F	: Banach spaces
$GL(E)$	: the Lie group of continuous automorphisms of E
$\mathfrak{gl}(E)$	: the Lie algebra of continuous endomorphisms of E
$L(E, F)$	: the set of continuous linear maps from E to F
$L^k(E, F)$	: the set of continuous k-multilinear maps from $E \times \cdots \times E$ to F
$\mathscr{L}_X$	: the Lie derivative with respect to $X \in \mathscr{T}(M)$
$\mathscr{T}(M)$	: the set of smooth vector fields on M
TM	: the tangent bundle of a manifold M
$T\phi\colon TM \to TN$	: the tangent of a mapping $\phi\colon M \to N$ of manifolds M and N
$\mathring{\bigcup}_{x\in X}$	: disjoint union over $x \in X$
$U \Subset X$	: U is an open subset of X

2. Affine and linear connections

In this paper we shall work in the category of C^∞ reflexive Banach manifolds. Here we present some notation for the basic concepts of affine and linear connections. Our geometric notation will follow that of [Abraham, Marsden, and Ratiu 1988]. For affine connections on Banach manifolds we refer the reader to [Lang 1995]. We caution that only torsion free connections are dealt with in Lang and here we shall need to allow connections with non-zero torsion. For linear connections we refer to [Kobayashi and Nomizu 1963] although this work is presented in the finite-dimensional context. As we shall see, it is straightforward to extend this to infinite-dimensions. As we will be interested in transformations of affine connections in Sections 3 and 5, we thoroughly review their basic properties here.

Affine connections. Let ∇ be an affine connection on a manifold M. If (U, ϕ) is a chart for M taking values in a Banach space E, we shall denote by $\Gamma\colon \phi(U) \to L^2(E, E)$ the ***Christoffel symbols*** for the affine connection in the chart. Thus in the chart we have[1]

$$\nabla_X Y(u) = (u, \boldsymbol{D}Y(u) \cdot X(u) + \Gamma(u)(X(u), Y(u))).$$

We shall denote by Z_g the geodesic spray of ∇ which is therefore a second-order vector field on TM and its integral curves project to geodesics of ∇. In a natural chart $(TU, T\phi)$ for TM we have $Z_g(u, e) = ((u, e), (e, -\Gamma(u)(e, e)))$. We denote the ***torsion*** and ***curvature*** tensors of ∇ by T and R and we recall that

$$\begin{aligned} T(X, Y) &= \nabla_X Y - \nabla_Y X - [X, Y] \\ R(X, Y)Z &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \end{aligned}$$

Thus we regard $R(X, Y)$ as a $(1, 1)$ tensor field for each pair of vector fields X and Y. Recall that if M has a Riemannian metric[2] then there is associated to it a unique affine connection ∇ with the properties that ∇ is torsion free and that $\nabla_X g = 0$ for every $X \in \mathscr{T}(M)$. This affine connection is called the ***Levi-Civita connection***.

Linear connections. Closely related to affine connections are linear connections which are principal connections on the bundle of linear frames. We suppose that M is modelled on a Banach space E and for $x \in M$ we denote by $L_x(M)$ the set of isomorphisms $p\colon E \to T_x M$. We write

$$L(M) = \mathring{\bigcup}_{x \in M} L_x(M)$$

which we call the ***bundle of linear frames*** of M even though this name is a bit misleading in the infinite-dimensional context. The group $GL(E)$ acts on the right on $L(M)$ by $(p, a) \mapsto p \circ a$ where $p \in L_x(M)$ for some $x \in M$ and $a \in GL(E)$. Let (U, ϕ) be a chart for M and let $x \in U$. The natural tangent bundle chart $(TU, T\phi)$ for TM provides an isomorphism of E with $T_x M$ via $(T_x \phi)^{-1}$. Given $p \in L_x(M)$ there exists a unique $a \in GL(E)$ such that $p = (T_x \phi)^{-1} \circ a$ and so this establishes a trivialisation $L(U) \simeq \phi(U) \times GL(E)$. Since it is clear that $L(M)/GL(E) \simeq M$, we are justified in saying that $(\pi, L(M), M, GL(E))$ is a principal fibre bundle with total space $L(M)$, base space M, structure group $GL(E)$, and projection $\pi\colon L(M) \to M$. A ***linear connection*** is a specification of a $GL(E)$-invariant

[1] One cannot actually write the covariant derivative in this form without making additional assumptions on the model Banach space. These assumptions are true in finite dimensions or, more generally, for Hilbert manifolds. We refer the reader to [Lang 1995, Chapter VIII, §2] for a discussion of these issues. Here we shall simply assume that the covariant derivative locally has the given form.

[2] We will assume strongly nondegenerate metrics unless we say otherwise. This guarantees the existence of the associated affine connection.

complement, $H(L(M))$, to the subbundle $V(L(M)) \triangleq \ker(T\pi)$. If $X \in T_p(L(M))$ then we will write $X = \mathrm{hor}(X) + \mathrm{ver}(X)$ where $\mathrm{hor}(X) \in H_p(L(M))$ and $\mathrm{ver}(X) \in V_p(L(M))$. The ***connection one-form*** ω is defined by

$$\omega_p(X) = \{A \in \mathfrak{gl}(E) \mid A^*(p) = \mathrm{ver}(X)\}$$

where A^* is the ***infinitesimal generator*** defined by

$$A^*(p) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (p \circ \exp(tA)).$$

Note that ω is a $\mathfrak{gl}(E)$-valued one-form on $L(M)$. Since $GL(E)$ acts freely on $L(M)$, ω is well-defined. On $L(M)$ there is also a canonically defined, E-valued one-form θ given by

$$\theta_p(X) = p^{-1}(T_p\pi(X)).$$

If $c\colon [a,b] \to M$ is a curve and if $p \in \pi^{-1}(c(a))$, we may define a curve $\sigma\colon [a,b] \to L(M)$ by horizontally lifting c through p. Thus $c = \pi \circ \sigma$ and $T\pi(\dot\sigma(t)) = \dot c(t)$ for $t \in [a,b]$. We define the ***parallel translation*** of p along c to be the section $t \mapsto \sigma(t)$ of $L(M)$ over c.

Note that $GL(E)$ acts on E on the left in the natural manner and so we have a left action of $GL(E)$ on $L(M)\times E$ given by $(a,(p,e)) \mapsto (p\circ a^{-1}, a(e))$. The quotient $(L(M)\times E)/GL(E)$ is a vector bundle associated with $L(M)$ and it is naturally isomorphic to TM. Parallel transport in $L(M)$ provides parallel transport in TM as follows. Let $c\colon [a,b] \to M$ be a curve with $v \in T_{c(a)}M$. Let $e \in E$ and define $p \in \pi^{-1}(c(a))$ by asking that $pe = v$. If σ is the horizontal lift of c through p then $t \mapsto \sigma(t)$ is the parallel translation of p along c. We define the parallel translation of v along c to be the vector field along c defined by $X(t) = \sigma(t)e$. One may show that this construction does not depend on $e \in E$. Associated with this parallel translation operation is an affine connection ∇ on M given by

$$(\nabla_X Y)(x) = p(\mathrm{hlft}_p(X)(\theta(\mathrm{hlft}_p(Y)))) \tag{2.1}$$

where $p \in \pi^{-1}(x)$ (see the Lemma on page 133 of [Kobayashi and Nomizu 1963]). Here $\mathrm{hlft}_p\colon T_xM \to H_p(L(M))$ is the natural horizontal lift given by the connection.

If $e \in E$, we denote by $\Sigma(e)$ the horizontal vector field on $L(M)$ defined by $T_p\pi(\Sigma(e)(p)) = pe \in T_{\pi(p)}M$. That is to say, $\Sigma(e)(p)$ is the horizontal lift of pe. One may show that $\sigma\colon [a,b] \to L(M)$ is an integral curve for $\Sigma(e)$ if and only if $\pi \circ \sigma$ is a geodesic for the associated affine connection ∇. Conversely, $c\colon [a,b] \to M$ is a geodesic for ∇ with initial condition $\dot c(a) = v_x \in T_xM$ if and only if its horizontal lift through $p \in L_x(M)$ is an integral curve for $\Sigma(p^{-1}v_x)$.

Transformations for affine connections. We refer the reader to [Kobayashi and Nomizu 1963, Chapter VI] for details on the presentation in this section. A diffeomorphism $\phi\colon M \to M$ is called an ***affine transformation*** of ∇ if any one of the following equivalent conditions is true:

AT1. $T\phi$ commutes with parallel translation;

AT2. $\phi \circ \exp_x = \exp_{\phi(x)} \circ T_x\phi$ for $x \in M$;

AT3. $\phi^*(\nabla_X Y) = \nabla_{\phi^* X}\phi^* Y$ for vector fields X and Y on M.

If ∇ is the Levi-Civita connection associated with a Riemannian metric g on M, then the above conditions are implied by

AT4. $\phi^* g = g$.

In this case we will sometimes say that ϕ is an ***isometry*** of the Riemannian metric g. We note that the set of affine transformations of ∇ forms a subgroup of the diffeomorphism group of M which we shall denote by $\mathit{Aff}(\nabla)$.

A vector field X is called an ***infinitesimal affine transformation*** if it satisfies any one of the following equivalent conditions:

IAT1. the flow of X consists of a one-parameter family of affine transformations;

IAT2. $\mathscr{L}_X \circ \nabla_Y - \nabla_Y \circ \mathscr{L}_X = \nabla_{[X,Y]}$ for every vector field Y.

If ∇ is the Levi-Civita connection associated with a Riemannian metric g then the above conditions are implied by:

IAT3. $\mathscr{L}_X g = 0$;

IAT4. $g(\nabla_Y X, Z) + g(Y, \nabla_Z X) = 0$ for vector fields Y and Z on M.

In this case we will sometimes say that X is an ***infinitesimal isometry*** of g or a ***Killing vector field***. One may show that the set of infinitesimal affine transformations of ∇ is a Lie subalgebra of $\mathscr{T}(M)$ which we shall denote by $\mathfrak{aff}(\nabla)$.

Since a proof of IAT4 is difficult to obtain in the literature, we present one here. We also add an additional helpful characterisation of Killing vector fields.

2.1 Lemma: *Let g be a Riemannian metric on M, let ∇ be the associated Levi-Civita connection, and let X be a vector field on M. The following are equivalent:*

(i) *X is an infinitesimal isometry for g;*

(ii) *$g(\nabla_Y X, Z) + g(Y, \nabla_Z X) = 0$ for every two vector fields Y and Z;*

(iii) *$g(\nabla_Y X, Y) = 0$ for every vector field Y.*

Proof: By continuity we may prove the lemma at those points $x \in M$ for which $X(x) \neq 0$. Let x be such a point and let (U, ϕ) be a chart around x which straightens out X. Thus for $x' \in U$, $\phi(x') = (t, u) \in V \times U'$ with $V \Subset \mathbb{R}$ and $U' \Subset E$. In this chart the local representative for X is $X(t, u) = ((t, u), (1, 0))$. Let us denote by $\Gamma \colon V \times U' \to L^2(\mathbb{R} \times E, \mathbb{R} \times E)$ the Christoffel symbols in this chart. Let $X_1, X_2 \in \mathbb{R} \times E$ and regard these as constant vector fields on $V \times U'$. Noting that X, X_1, and X_2 are constant we compute

$$
\begin{aligned}
g(\nabla_{X_1} X, X_2) &+ g(X_1, \nabla_{X_2} X) \\
&= g(\boldsymbol{D}X \cdot X_1 + \Gamma(X_1, X), X_2) + g(X_1, \boldsymbol{D}X \cdot X_2 + \Gamma(X_2, X)) \\
&= g(\boldsymbol{D}X_1 \cdot X + \Gamma(X, X_1), X_2) + g(X_1, \boldsymbol{D}X_2 \cdot X + \Gamma(X, X_2)) \\
&= g(\nabla_X X_1, X_2) + g(X_1, \nabla_X X_2). \qquad (2.2)
\end{aligned}
$$

Here we have used the fact that the Levi-Civita connection has zero torsion and so Γ is symmetric.

(i) $\iff$ (ii) Now we compute

$$\begin{aligned}\mathscr{L}_X(g(X_1,X_2)) &= (\mathscr{L}_X g)(X_1,X_2) + g(\mathscr{L}_X X_1, X_2) + g(X_1, \mathscr{L}_X X_2)\\ &= (\mathscr{L}_X g)(X_1,X_2)\end{aligned}$$

since X commutes with X_1 and X_2. We also have

$$\begin{aligned}\mathscr{L}_X(g(X_1,X_2)) &= (\nabla_X g)(X_1,X_2) + g(\nabla_X X_1, X_2) + g(X_1, \nabla_X X_2)\\ &= g(\nabla_X X_1, X_2) + g(X_1, \nabla_X X_2)\end{aligned}$$

since the Levi-Civita connection has the property that $\nabla_Z g = 0$ for every vector field Z. Therefore, by (2.2)

$$g(\nabla_{X_1} X, X_2) + g(X_1, \nabla_{X_2} X) = g(\nabla_X X_1, X_2) + g(X_1, \nabla_X X_2) = (\mathscr{L}_X g)(X_1, X_2).$$

Thus we see that X is an infinitesimal affine transformation for ∇ if and only if $g(\nabla_{X_1} X, X_2) + g(X_1, \nabla_{X_2} X) = 0$ for every $X_1, X_2 \in \mathbb{R} \times E$. Since this expression is linear with respect to multiplication by functions in X_1 and X_2 and since X_1 and X_2 are arbitrary, we see that (i) and (ii) are equivalent.

(i) $\iff$ (iii) We claim that for any vector field X, $\mathscr{L}_X g$ is a symmetric $(0,2)$ tensor field. Indeed we have

$$(\mathscr{L}_X g)(Y,Z) = \mathscr{L}_X(g(Y,Z)) - g(\mathscr{L}_X Y, Z) - g(Y, \mathscr{L}_X Z)$$

which is clearly symmetric in Y and Z. Therefore, $\mathscr{L}_X g = 0$ if and only if $\mathscr{L}_X g(Y,Y) = 0$ for every vector field Y. Letting $Y \in \mathbb{R} \times E$ we have

$$\mathscr{L}_X g(Y,Y) = 2g(\nabla_X Y, Y) = 2g(\nabla_Y X, Y)$$

by (2.2). Thus we see that X is a Killing vector field if and only if $g(\nabla_Y X, Y) = 0$ for every $Y \in \mathbb{R} \times E$. Again, this expression is linear with respect to multiplication by a function in the argument Y, so we see that X is a Killing vector field if and only if (iii) holds. ■

Motivated by property IAT2 of infinitesimal affine transformations, let us define

$$B_X(Y,Z) = [X, \nabla_Y Z] - \nabla_Y [X,Z] - \nabla_{[X,Y]} Z. \tag{2.3}$$

We claim that B_X is a tensor field of type $(1,2)$ for every $X \in \mathscr{T}(M)$. To see this, let f and g be functions on M and compute

$$\begin{aligned}B_X(fY, gZ) &= [X, \nabla_{fY} gZ] - \nabla_{fY}[X, gZ] - \nabla_{[X,fY]}(gZ)\\ &= [X, fg\nabla_Y Z + f(\mathscr{L}_Y g)Z] - f\nabla_Y(g[X,Z] + (\mathscr{L}_X g)Z) -\\ &\quad g\nabla_{f[X,Y]+(\mathscr{L}_X f)Y} Z - (\mathscr{L}_{f[X,Y]+(\mathscr{L}_X f)Y} g)Z\\ &= fg[X, \nabla_Y Z] + (\mathscr{L}_X fg)\nabla_Y Z + f(\mathscr{L}_Y g)[X,Z] + \mathscr{L}_X(f(\mathscr{L}_Y g))Z -\\ &\quad fg\nabla_Y[X,Z] - f(\mathscr{L}_Y g)[X,Z] - f(\mathscr{L}_X g)\nabla_Y Z - f\mathscr{L}_Y(\mathscr{L}_X g)Z -\\ &\quad fg\nabla_{[X,Y]} Z - g(\mathscr{L}_X f)\nabla_Y Z - f(\mathscr{L}_{[X,Y]} g)Z - (\mathscr{L}_X f)(\mathscr{L}_Y g)Z\\ &= fg[X, \nabla_Y Z] - fg\nabla_Y[X,Z] - fg\nabla_{[X,Y]} Z = fgB_X(Y,Z).\end{aligned}$$

This verifies that B_X is a tensor field. The following lemma follows from the definition of B_X.

2.2 Lemma: *A vector field X is an infinitesimal affine transformation for ∇ if and only if $B_X = 0$.*

3. Affine connections which restrict to a distribution

An affine connection ∇ on M is said to ***restrict*** to a distribution D if $\nabla_X Y \in \mathscr{D}$ for every $Y \in \mathscr{D}$. Let $v \in T_xM$ and let $c\colon [0,T] \to M$ be a curve such that $c(0) = x$. Recall that the ***parallel translation*** of v along c is the vector field X along c which satisfies the differential equation $\nabla_{\dot{c}(t)} X(t) = 0$ with initial condition $X(0) = v$. Thus we see that if ∇ restricts to D, then parallel translation leaves D invariant. In this section we study some general properties of affine connections which restrict to a distribution.

> *Unless stated otherwise, throughout the remainder of this section, let ∇ be an affine connection which restricts to D.*

3.1. Torsion, curvature, and holonomy.

The property of an affine connection restricting to a distribution makes it possible to simplify the torsion and curvature of the connection. We note in particular how it is important that one allow non-zero torsion for affine connections which restrict to certain distributions. Let T and R denote the torsion and curvature tensors, respectively.

3.1 Proposition: *D is integrable if and only if $T(X,Y) \in \mathscr{D}$ for every $X, Y \in \mathscr{D}$. In particular, if $T = 0$ then D is integrable.*

Proof: This follows from the definition of T and our assumption that ∇ restricts to D. ■

Another way of stating the above result is to say that D is integrable if and only if the torsion of ∇ restricts to D. The above result has the following interesting corollary.

3.2 Corollary: *Let (M, g) be a Riemannian manifold with ∇ the associated Levi-Civita affine connection. Then ∇ restricts to a distribution D only if D is integrable.*

The "if" direction of the corollary is obviously not true in general.

3.3 Proposition: *Let $x \in M$ and let $u, v \in T_xM$. Then the endomorphism $R(u,v)$ of T_xM leaves invariant the subspace D_x.*

Proof: This follows from the definition of the curvature tensor and from the fact that we are supposing ∇ restricts to D. ■

In a related manner, we may make an essential observation about the holonomy groups of ∇. The reader will wish to recall the discussion of holonomy groups for linear connections in Section III.9 of [Kobayashi and Nomizu 1963].[3] In particular, recall that the holonomy group at $x \in M$ may be thought of as a subgroup of the automorphism group of T_xM. The Lie algebra of this group is then a subalgebra of the Lie algebra of endomorphisms of T_xM equipped with the commutator bracket.

---

[3] One may easily convince oneself that the results presented by Kobayashi and Nomizu extend to the infinite-dimensional setting.

3.4 Proposition: *Suppose that ∇ restricts to D and let $x \in M$. The following statements hold.*

(i) D_x is an invariant subspace of the holonomy group of ∇ at x.

(ii) The Lie algebra of the infinitesimal holonomy group of ∇ at x has D_x as an invariant subspace.

Proof: (i) The holonomy group of ∇ at x is generated by the automorphisms

$$T_xM \ni w \mapsto \tau_c^{-1} \circ R(\tau_c(u), \tau_c(v)) \circ \tau_c(w)$$

where $c\colon [0,1] \to M$ is a curve with $c(0) = x$, τ_c is parallel translation along c, and $u, v \in T_xM$. Since ∇ restricts to D, D is invariant under parallel translation. By Proposition 3.3 $R(\tau_c(u), \tau_c(v)) \circ \tau_c(w) \in D_{c(1)}$. This proves (i).

(ii) The Lie algebra of the infinitesimal holonomy group at x is generated by endomorphisms of T_xM of the form

$$\nu \mapsto (\nabla^k R)(u, v; w_1; \dots; w_k)(\nu)$$

where $u, v, w_1, \dots, w_k \in T_xM$ and $k \in \mathbb{N}$. We shall prove by induction on k that each endomorphism in this generating set leaves D_x invariant. This is true for $k = 0$ by Proposition 3.3. Now suppose it is true for $k = l > 0$. Therefore, for vector fields $X, Y, V_1, \dots, V_l$ on M and for any section Z of D,

$$(\nabla^l R)(X, Y, Z; V_1; \dots; V_l) \triangleq (\nabla^l R)(X, Y; V_1; \dots; V_l)(Z) \in \mathscr{D}.$$

If V_{l+1} is a vector field on M we compute

$$\begin{aligned}(\nabla^{l+1} R)(X, Y, Z; &V_1; \dots; V_l; V_{l+1}) = \left(\nabla_{V_{l+1}}(\nabla^l R)\right)(X, Y, Z; V_1; \dots; V_l)\\ &= \nabla_{V_{l+1}}\left(\nabla^l R(X, Y, Z; V_1; \dots; V_l)\right) - (\nabla^l R)(\nabla_{V_{l+1}} X, Y, Z; V_1; \dots; V_l) -\\ &\quad (\nabla^l R)(X, \nabla_{V_{l+1}} Y, Z; V_1; \dots; V_l) - (\nabla^l R)(X, Y, \nabla_{V_{l+1}} Z; V_1; \dots; V_l) -\\ &\quad \sum_{i=1}^{l} (\nabla^l R)(X, Y, Z; V_1; \dots; \nabla_{V_{l+1}} V_i; \dots; V_l).\end{aligned}$$

By the induction hypothesis and since we are supposing that ∇ restricts to D, each of the terms on the right hand side of the final expression is a section of D. This completes the proof of (ii). ■

3.2. Transformations. In this section we discuss transformations of affine connections which restrict to a distribution. We look for transformations which respect the affine connection *and* the distribution. We shall say that a map $\phi\colon M \to M$ is ***compatible with D*** if $T_x\phi(D_x) = D_{\phi(x)}$ for each $x \in M$. A vector field X is ***compatible with D*** if $[X, Y] \in \mathscr{D}$ for every $Y \in \mathscr{D}$. It is common to use the expression "D is invariant under X" if X is compatible with D.

Setting aside issues of completeness of vector fields, we have the following lemma.

3.5 Lemma: *A vector field X is compatible with D if and only if its flow is compatible with D.*

Proof: Denote by F_t the flow of X and suppose that F_t is compatible with D. If $Y \in \mathscr{D}$ we have

$$\mathscr{L}_X Y(x) = \lim_{t\to 0} \left(T_{F_t(x)}F_{-t}(Y(F_t(x))) - Y(x)\right).$$

Since F_t is compatible with D, $T_{F_t(x)}F_{-t}(Y(F_t(x))) \in D_x$ and therefore $\mathscr{L}_X Y(x) \in D_x$ and so X is compatible with D.

To prove the converse we make the following observation. Let X^T denote the lift of X to TM. Thus X^T is the vector field on TM whose flow is the one-parameter family of diffeomorphisms $t \mapsto TF_t$. We claim that F_t is compatible with D if and only if $D \subset TM$ is invariant under the flow of X^T. Indeed, F_t is compatible with D if and only if $T_xF_t(D_x) = D_{F_t(x)}$ for each $x \in M$ from which our claim easily follows. Thus we shall prove that X^T is tangent to D if and only if X is compatible with D and this will complete the proof of the lemma. We shall use coordinates for TM which are adapted to the distribution. Thus we choose a chart (U, ϕ) for M with $(TU, T\phi)$ the induced natural chart for TM. Since D is a subbundle of TM, there exists a chart (TU, ψ) such that

1. ψ is a bijection onto $U' \times F_1 \times F_2$ where $U' = \phi(U) \Subset E$, and F_1 and F_2 are Banach spaces,
2. $\psi(D_{\phi^{-1}(u)}) = \{u\} \times F_1 \times \{0\}$,
3. the overlap map from $T\phi(TU)$ to $\psi(TU)$ has the form
$$h\colon (u, e) \mapsto (u, A_1(u)\cdot e, A_2(u)\cdot e)$$
for smooth maps $A_i\colon U' \to L(E, F_i)$, $i = 1, 2$.

We denote the inverse of h by

$$h^{-1}\colon (u, f_1, f_2) \mapsto (u, B_1(u)\cdot f_1 + B_2(u)\cdot f_2)$$

which defines smooth maps $B_i\colon U' \to L(F_i, E)$, $i = 1, 2$.

Let $(u, e) \in U' \times E$ and let $(u, f_1, f_2) = h(u, e) \in U' \times F_1 \times F_2$. We compute

$$\begin{aligned}
\boldsymbol{D}h(u,e)\cdot X^T(u,e) &= \boldsymbol{D}(u,e)\cdot(X(u), \boldsymbol{D}X(u)\cdot e)\\
&= (X(u), \boldsymbol{D}A_1(u)(X(u), e) + A_1(u)\cdot(\boldsymbol{D}X(u)\cdot e),\\
&\quad \boldsymbol{D}A_2(u)(X(u), e) + A_2(u)\cdot(\boldsymbol{D}X(u)\cdot e))\\
&= (X(u), \boldsymbol{D}A_1(u)(X(u), B_1(u)\cdot f_1 + B_2(u)\cdot f_2) +\\
&\quad A_1(u)\cdot(\boldsymbol{D}X(u)\cdot(B_1(u)\cdot f_1 + B_2(u)\cdot f_2)),\\
&\quad \boldsymbol{D}A_2(u)(X(u), B_1(u)\cdot f_1 + B_2(u)\cdot f_2) +\\
&\quad A_2(u)\cdot(\boldsymbol{D}X(u)\cdot(B_1(u)\cdot f_1 + B_2(u)\cdot f_2))).
\end{aligned}$$

If we restrict this representation of X^T to $D \subset TM$ by setting $f_2 = 0$ we get

$$\begin{aligned}
T\psi(X^T|D) = (X(u), &\boldsymbol{D}A_1(u)(X(u), B_1(u)\cdot f_1) + A_1(u)\cdot(\boldsymbol{D}X(u)\cdot(B_1(u)\cdot f_1)),\\
&\boldsymbol{D}A_2(u)(X(u), B_1(u)\cdot f_1) + A_2(u)\cdot(\boldsymbol{D}X(u)\cdot(B_1(u)\cdot f_1))). \quad (3.1)
\end{aligned}$$

Now let $f_1 \in F_1$ and consider the vector field $Y\colon u \mapsto (u, B_1(u)\cdot f_1)$ on U' which is a representative of a section of D. We compute

$$\begin{aligned} h([X,Y](u)) &= h(u, \boldsymbol{D}B_1(u)(X(u),f_1) - \boldsymbol{D}X(u)\cdot(B_1(u)\cdot f_1)) \\ &= (u, A_1(u)\cdot(\boldsymbol{D}B_1(u)(X(u),f_1) - \boldsymbol{D}X(u)\cdot(B_1(u)\cdot f_1)), \\ &\quad A_2(u)\cdot(\boldsymbol{D}B_1(u)(X(u),f_1) - \boldsymbol{D}X(u)\cdot(B_1(u)\cdot f_1))). \end{aligned} \tag{3.2}$$

By definition of A_2 and B_1 we have $A_2(u)\cdot(B_1(u)\cdot f_1) = 0$ for every $f_1 \in F_1$. Differentiating this in the direction $X(u)$ we get

$$\boldsymbol{D}A_2(u)(X(u), B_1(u)\cdot f_1) + A_2(u)\cdot(\boldsymbol{D}B_1(X(u),f_1)) = 0. \tag{3.3}$$

Combining (3.1), (3.2), and (3.3) we obtain the lemma. ■

We now wish to consider an affine connection ∇ which restricts to D. Recall the definition of B_X from (2.3).

3.6 Lemma: *Let X be a vector field compatible with D and suppose that ∇ restricts to D. Then $B_X(Y,Z) \in \mathscr{D}$ for every $Y, Z \in \mathscr{D}$.*

Proof: Let $Y, Z \in \mathscr{D}$. We have

$$B_X(Y,Z) = [X, \nabla_Y Z] - \nabla_Y [X,Z] - \nabla_{[X,Y]} Z.$$

Since ∇ restricts to D, $\nabla_Y Z \in \mathscr{D}$ and since X is compatible with D, $[X, \nabla_Y Z] \in \mathscr{D}$. In similar fashion we see that $\nabla_Y[X,Z] \in \mathscr{D}$ and $\nabla_{[X,Y]}Z \in \mathscr{D}$ thus proving the lemma. ■

This indicates that, under the hypotheses of the lemma, one may restrict considerations of infinitesimal affine invariance to the vector bundle connection in D. This leads to the following definitions.

3.7 Definition: Let $\phi\colon M \to M$ be a diffeomorphism which is compatible with D, let X be a vector field which is compatible with D, and suppose that ∇ restricts to D. We say that ϕ is a ***D-affine transformation*** if $T\phi$ commutes with parallel translation in D, and we say that X is a ***D-infinitesimal affine transformation*** if its flow comprises a one-parameter family of D-affine transformations. •

Since ∇ restricts to D, the above definitions make sense. The following result is simply proved by restricting the statements for general affine connections to the distribution D. Everything goes through since we are supposing ∇ leaves D invariant.

3.8 Proposition: *Suppose that ∇ restricts to D and let ϕ be a diffeomorphism compatible with D. The following are equivalent:*

(i) ϕ *is a D-affine transformation;*

(ii) $\phi \circ (\exp_x |D_x) = \exp_{\phi(x)} \circ (T_x\phi | D_x)$;

(iii) $\phi^*(\nabla_X Y) = \nabla_{\phi^* X}\phi^* Y$ *if* $Y \in \mathscr{D}$.

If ∇ is the Levi-Civita connection associated with a Riemannian metric g, then the above conditions are implied by:

(iv) $(\phi^* g)(X,Y) = g(X,Y)$ *for* $X, Y \in \mathscr{D}$.

Let X be a vector field compatible with D. The following are equivalent:

(v) X *is a* D*-infinitesimal affine transformation;*

(vi) $B_X(Y,Z)=0$ *for* $Y,Z\in\mathscr{D}$.

If ∇ *is the Levi-Civita connection associated with a Riemannian metric* g*, then the above conditions are implied by:*

(vii) $(\mathscr{L}_X g)(Y,Z)=0$ *for* $Y,Z\in\mathscr{D}$;

(viii) $g(\nabla_Y X,Z)+g(Y,\nabla_Z X)=0$ *for* $Y,Z\in\mathscr{D}$;

(ix) $g(\nabla_Y X,Y)=0$ *for* $Y\in\mathscr{D}$.

Next we consider the properties of D-affine transformations under composition and D-infinitesimal affine transformations under Lie bracket. Let us denote by $\mathit{Diff}(M)$ the group of diffeomorphisms of M, by $\mathit{Aff}(\nabla|D)$ the set of D-affine transformations, and by $\mathfrak{aff}(\nabla|D)$ the set of D-infinitesimal affine transformations.

3.9 Proposition: *Suppose that* ∇ *restricts to* D*. Then*

(i) $\mathit{Aff}(\nabla|D)$ *is a subgroup of* $\mathit{Diff}(M)$,

(ii) $\mathfrak{aff}(\nabla|D)$ *is a Lie subalgebra of* $\mathscr{T}(M)$.

Proof: (i) First note that if diffeomorphisms ϕ_1 and ϕ_1 are compatible with D, then $\phi_1\circ\phi_2$ is compatible with D. For $Y\in\mathscr{D}$ and $\phi_1,\phi_2\in\mathit{Aff}(\nabla|D)$ we compute

$$\begin{aligned}(\phi_1\circ\phi_2)^*(\nabla_X Y)&=\phi_2^*(\phi_1^*\nabla_X Y)\\&=\phi_2^*\nabla_{\phi_1^*X}\phi_1^*Y\\&=\nabla_{(\phi_1\circ\phi_2)^*X}(\phi_1\circ\phi_2)^*Y\end{aligned}$$

by Proposition 3.8(iii). By that same result we then see that $\phi_1\circ\phi_2\in\mathit{Aff}(\nabla|D)$. Also note that $\mathrm{id}_M\in\mathit{Aff}(\nabla|D)$ which shows that $\mathit{Aff}(\nabla|D)$ is a subgroup of $\mathit{Diff}(M)$.

(ii) First we claim that if $X,Y\in\mathscr{T}(M)$ are compatible with D, then $[X,Y]$ is compatible with D. Indeed, for $Z\in\mathscr{D}$ we have

$$[[X,Y],Z]=-[[Z,X],Y]-[[Y,Z],X]$$

by Jacobi's identity. From this our claim follows. Now let $X,Y\in\mathfrak{aff}(\nabla|D)$. Then we have

$$(\mathscr{L}_X\circ\nabla_Z-\nabla_Z\circ\mathscr{L}_X)W=(\nabla_{[X,Z]})W,\quad(\mathscr{L}_Y\circ\nabla_Z-\nabla_Z\circ\mathscr{L}_Y)W=(\nabla_{[Y,Z]})W$$

for every $Z,W\in\mathscr{D}$ by Proposition 3.8(vi). We now compute

$$\begin{aligned}(\mathscr{L}_{[X,Y]}\circ\nabla_Z-&\nabla_Z\circ\mathscr{L}_{[X,Y]})W=\\&(\mathscr{L}_X\circ\mathscr{L}_Y\circ\nabla_Z-\mathscr{L}_Y\circ\mathscr{L}_X\nabla_Z-\nabla_Z\circ\mathscr{L}_X\circ\mathscr{L}_Y+\nabla_Z\circ\mathscr{L}_Y\circ\mathscr{L}_X)W\\&=(\mathscr{L}_X\circ(\mathscr{L}_Y\circ\nabla_Z-\nabla_Z\circ\mathscr{L}_Y)+(\mathscr{L}_X\nabla_Z-\nabla_Z\circ\mathscr{L}_X)\circ\mathscr{L}_Y-\\&\quad\mathscr{L}_Y\circ(\mathscr{L}_X\nabla_Z-\nabla_Z\circ\mathscr{L}_X)-(\mathscr{L}_Y\circ\nabla_Z-\nabla_Z\circ\mathscr{L}_Y)\circ\mathscr{L}_X)W\\&=(\mathscr{L}_X\circ\nabla_{[Y,Z]}-\nabla_{[Y,Z]}\circ\mathscr{L}_X-\mathscr{L}_Y\circ\nabla_{[X,Z]}+\nabla_{[X,Z]}\circ\mathscr{L}_Y)W\\&=(\nabla_{[X,[Y,Z]]}-\nabla_{[Y,[X,Z]]})W\\&=(\nabla_{[[X,Y],Z]})W.\end{aligned}$$

The penultimate equality comes from Proposition 3.8(vi) (noting that $[X,Z],[Y,Z]\in\mathscr{D}$) and the final equality follows from Jacobi's identity. We thus see that $B_{[X,Y]}(Z,W)=0$ for every $Z,W\in\mathscr{D}$ and so $[X,Y]\in\mathfrak{aff}(\nabla|D)$. ■

It is not altogether clear what relationship exists between $\mathit{Aff}(\nabla)$ and $\mathit{Aff}(\nabla|D)$ (or their respective Lie algebras). In Section 5 (cf. Proposition 5.7) we see that in at least one case they have some transformations in common. Also, in the finite-dimensional case at least, one should be able to demonstrate that $\mathit{Aff}(\nabla|D)$ is a finite-dimensional Lie group with Lie algebra $\mathfrak{aff}(\nabla|D)$.

Finally we consider the case when D has a complement D'. We shall say a diffeomorphism $\phi\colon M \to M$ (resp. a vector field X on M) is ***compatible*** with (D, D') if ϕ is compatible with D and D' (resp. X is compatible with D and D'). This arises in one important case as the next lemma illustrates. In this lemma we do not assume that ∇ restricts to D.

3.10 Lemma: *Suppose that M comes equipped with a Riemannian metric g and that D is a distribution with $D^\perp$ its orthogonal complement. Then*

(i) an isometry, ϕ, for the Levi-Civita connection is compatible with D if and only if it is compatible with $(D, D^\perp)$,

(ii) an infinitesimal isometry, X, for the Levi-Civita connection is compatible with D if and only if it is compatible with $(D, D^\perp)$.

Proof: (i) It is clear that if ϕ is compatible with $(D, D^\perp)$ then it is compatible with D. Now suppose that ϕ is compatible with D, and let $Y \in \mathscr{D}$ and $Z \in \mathscr{D}^\perp$. Since ϕ is compatible with D, $\phi_* Y \in \mathscr{D}$. Therefore

$$\begin{aligned} & g(\phi_* Y, Z) = 0 \\ \implies\quad & \phi^*(g(\phi_* Y, Z)) = 0 \\ \implies\quad & (\phi^* g)(Y, \phi^* Z) = 0 \\ \implies\quad & g(Y, \phi^* Z) = 0. \end{aligned}$$

Here we have used the fact that ϕ is an isometry. Since Y is arbitrary, $\phi^* Z \in \mathscr{D}^\perp$, proving that ϕ is compatible with $D^\perp$.

(ii) It is clear that if X is compatible with $(D, D^\perp)$ then it is compatible with D. Thus we only show the converse. Suppose that $Y \in \mathscr{D}$, $Z \in \mathscr{D}^\perp$, and that X is compatible with D. Then

$$\begin{aligned} & g(Y, Z) = 0 \\ \implies\quad & \mathscr{L}_X(g(Y, Z)) = 0 \\ \implies\quad & (\mathscr{L}_X g)(Y, Z) + g(\mathscr{L}_X Y, Z) + g(Y, \mathscr{L}_X Z) = 0 \\ \implies\quad & g(Y, \mathscr{L}_X Z) = 0. \end{aligned}$$

Here we have used the fact that X is compatible with D and the fact that X is an infinitesimal affine transformation for the Levi-Civita connection. We then see that $\mathscr{L}_X Z \in \mathscr{D}^\perp$ since Y is arbitrary. ■

4. Geodesic invariance

In this section we generalise the notion of a totally geodesic submanifold.

4.1 Definition: A distribution D on a manifold M with an affine connection ∇ is ***geodesically invariant*** if for every geodesic $c\colon [a,b] \to M$ of ∇, $\dot{c}(a) \in D_{c(a)}$ implies that $\dot{c}(t) \in D_{c(t)}$ for every $t \in]a,b]$. •

It is clear that D is geodesically invariant if and only if Z_g is tangent to $D \subset TM$. Note that if D is a regular integrable distribution which is geodesically invariant, then each of its maximal integral manifolds is totally geodesic.

4.1. Characterising geodesically invariant distributions. To characterise geodesically invariant distributions, we shall use the ***symmetric product*** which is a product on $\mathscr{T}(M)$. For $X, Y \in \mathscr{T}(M)$ we define their symmetric product to be the vector field

$$\langle X : Y\rangle = \nabla_X Y + \nabla_Y X.$$

Given a vector field X on M it is possible to lift this to a vector field $\mathrm{vlft}(X)$ on TM which is tangent to the fibres. The vector field $\mathrm{vlft}(X)$ is defined by

$$\mathrm{vlft}(X)(v_x) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (v_x + tX(x))$$

for $v_x \in T_xM$. If (U, ϕ) is a chart for M with $(TU, T\phi)$ the associated natural chart for TM, then $\mathrm{vlft}(X)(u,e) = ((u,e),(0,X(u)))$. By definition of the vertical lift we have the following result.

4.2 Lemma: *Let D be a distribution on M. A vector field X on M is a section of D if and only if* $\mathrm{vlft}(X)$ *is tangent to the distribution D thought of as a submanifold of TM.*

The following formula for the vertical lift of the symmetric product will be useful.

4.3 Lemma: *Let X and Y be vector fields on M. Then*

$$\mathrm{vlft}(\langle X : Y\rangle) = [\mathrm{vlft}(X), [Z_g, \mathrm{vlft}(Y)]].$$

Proof: We use a chart (U, ϕ) for M with $(TU, T\phi)$ the associated natural chart for TM. It is an easy matter to compute

$$[Z_g, \mathrm{vlft}(Y)](u,e) = ((u,e),(-Y(u), \boldsymbol{D}Y(u)\cdot e + \Gamma(u)(Y(u),e))).$$

We then compute

$$\begin{aligned}&[\mathrm{vlft}(X), [Z_g, \mathrm{vlft}(Y)]](u,e) = \\ &\quad ((u,e),(0, \boldsymbol{D}X(u)\cdot Y(u) + \boldsymbol{D}Y(u)\cdot X(u) + \Gamma(u)(X(u),Y(u)) + \Gamma(u)(Y(u),X(u))))\end{aligned}$$

which we recognise as the representative of the vertical lift of $\langle X : Y\rangle$. ■

The following result provides infinitesimal tests for geodesic invariance and gives the geometric meaning of the symmetric product.

4.4 Theorem: *Let D be a distribution on a manifold M with an affine connection ∇. The following are equivalent:*

(i) D is geodesically invariant;

(ii) $\langle X : Y\rangle \in \mathscr{D}$ for every $X, Y \in \mathscr{D}$;

(iii) $\nabla_X X \in \mathscr{D}$ for every $X \in \mathscr{D}$.

Note that for geodesically invariant distributions the symmetric product plays the rôle that the Lie bracket plays for integrable distributions.

Proof of Theorem 4.4: (i) $\Longrightarrow$ (ii) First suppose that D is geodesically invariant. Thus Z_g is tangent to $D \subset TM$. Now let $X, Y \in \mathscr{D}$. By Lemma 4.2 $\mathrm{vlft}(X)$ and $\mathrm{vlft}(Y)$ are tangent to D. Therefore $[\mathrm{vlft}(X), [Z_g, \mathrm{vlft}(Y)]]$ is also tangent to D. By Lemmas 4.2 and 4.3 we see that $\langle X : Y\rangle \in \mathscr{D}$.

That (ii) implies (iii) follows from the definition of the symmetric product.

(iii) $\Longrightarrow$ (i) We work locally. Let (U, ϕ) be a chart for M taking values in E and denote by $(TU, T\phi)$ the associated natural chart for TM. Since D is a subbundle of TM we may choose a chart (TU, ψ) which has the properties discussed in the proof of Lemma 3.5. We borrow the notation used in that proof. Let $f_1 \in F_1$ and define a vector field on U' by $X(u) = (u, B_1(u) \cdot f_1)$. Note that X is a representative of a vector field taking values in D. We have

$$\begin{aligned} h(\nabla_X X(u)) &= h(u, \boldsymbol{D}B_1(u)(B_1(u)\cdot f_1, f_1) + \Gamma(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1)) \\ &= (u, A_1(u)\cdot(\boldsymbol{D}B_1(u)(B_1(u)\cdot f_1, f_1) + \Gamma(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1)), \\ &\quad\ A_2(u)\cdot(\boldsymbol{D}B_1(u)(B_1(u)\cdot f_1, f_1) + \Gamma(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1))). \end{aligned}$$

Since we are assuming (iii) we have

$$A_2(u)\cdot(\boldsymbol{D}B_1(u)(B_1(u)\cdot f_1, f_1) + \Gamma(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1))) = 0 \tag{4.1}$$

for every $f_1 \in F_1$. We now compute the geodesic spray in the chart (TU, ψ). Let $(u, e) \in U' \times E$ and let $(u, f_1, f_2) = h(u, e)$. We have

$$\begin{aligned} \boldsymbol{D}h(u,e)\cdot(e, -\Gamma(u)(e,e)) &= (e, \boldsymbol{D}A_1(u)(e,e) - A_1(u)\cdot\Gamma(u)(e,e), \\ &\qquad \boldsymbol{D}A_2(u)(e,e) - A_2(u)\cdot\Gamma(u)(e,e)) \\ &= (B_1(u)\cdot f_1 + B_2(u)\cdot f_2, \\ &\qquad \boldsymbol{D}A_1(u)(B_1(u)\cdot f_1 + B_2(u)\cdot f_2, B_1(u)\cdot f_1 + B_2(u)\cdot f_2) - \\ &\qquad A_1(u)\cdot\Gamma(u)(B_1(u)\cdot f_1 + B_2(u)\cdot f_2, B_1(u)\cdot f_1 + B_2(u)\cdot f_2), \\ &\qquad \boldsymbol{D}A_2(u)(B_1(u)\cdot f_1 + B_2(u)\cdot f_2, B_1(u)\cdot f_1 + B_2(u)\cdot f_2) - \\ &\qquad A_2(u)\cdot\Gamma(u)(B_1(u)\cdot f_1 + B_2(u)\cdot f_2, B_1(u)\cdot f_1 + B_2(u)\cdot f_2)). \end{aligned}$$

If we restrict this to D by setting $f_2 = 0$ we get

$$\begin{aligned} T\psi(Z_g|D) = (&B_1(u)\cdot f_1, \\ &\boldsymbol{D}A_1(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1) - A_1(u)\cdot\Gamma(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1), \\ &\boldsymbol{D}A_2(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1) - A_2(u)\cdot\Gamma(u)(B_1(u)\cdot f_1, B_1(u)\cdot f_1)). \end{aligned} \tag{4.2}$$

Note that by definition of A_2 and B_1 we have $A_2(u) \cdot (B_1(u) \cdot f_1) = 0$ for every $f_1 \in F_1$. If we differentiate this in the direction of $B_1(u) \cdot f_1$ we obtain

$$\boldsymbol{D}A_2(u)(B_1(u) \cdot f_1, B_1(u) \cdot f_1) + A_2(u) \cdot (\boldsymbol{D}B_1(u)(B_1(u) \cdot f_1, f_1)) = 0.$$

Combining this relation with (4.1) we see that the last component of (4.2) vanishes and so this shows that $Z_g|D$ is tangent to D. Thus D is geodesically invariant. ■

Note that if ∇ restricts to D, then D is geodesically invariant by (iii) of the theorem. The converse is not in general true.

To conclude our discussion in this section we add the following result which gives another interpretation of the symmetric product, at least in a chart.

4.5 Lemma: *Let ∇ be an affine connection on M with $\langle \cdot : \cdot \rangle$ the associated symmetric product. Let X and Y be vector fields on M with flows F_t^X and F_t^Y, respectively. Fix $x \in M$ and consider the following construction:*

(i) flow along the integral curve for X through x for time ϵ, arriving at x_ϵ;

(ii) flow along the integral curve for Y through x_ϵ for time ϵ, arriving at $x_{2\epsilon}$;

(iii) parallel transport $X(x_{2\epsilon})$ along the integral curve for Y to x_ϵ to get $Z_\epsilon \in T_{x_\epsilon}M$;

(iv) parallel transport Z_ϵ along the integral curve for X to get $Z_{2\epsilon} \in T_xM$.

Then

$$\langle X : Y \rangle(x) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} Z_{2\epsilon}.$$

Proof: If $c\colon I \to M$ is an integral curve for X and if $t_1, t_2 \in$I then let $\tau^X_{t_1,t_2}\colon T_{c(t_1)}M \to T_{c(t_2)}M$ denotes the parallel transport map. Along the integral curve of X from x to x_ϵ define a vector field $Z_1(t) = \tau^X_{\epsilon,t}(Z_\epsilon)$, $t \in [0, \epsilon]$. Similarly, along the integral curve of Y from x_ϵ to $x_{2\epsilon}$ define a vector field $Z_2(t) = \tau^Y_{\epsilon,t}(Z_{2\epsilon})$, $t \in [0, \epsilon]$. By the definition of covariant derivative we have

$$\nabla_Y X(x_\epsilon) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \tau^Y_\epsilon(X(x_{2\epsilon})),$$

so that

$$Z_\epsilon = \tau^Y_\epsilon(X(x_{2\epsilon})) = X(x_\epsilon) + \epsilon\nabla_Y X(x_\epsilon) + O(\epsilon^2).$$

■

4.2. Geodesic invariance and restricted affine connections. We have seen that the notion of geodesic invariance of a distribution is stronger than that of restriction of an affine connection to a distribution. In this section we give conditions under which a geodesically invariant distribution also has the property that the affine connection restricts to the distribution.

Let D be a distribution on M and suppose that M is modelled on a Banach space E. We say that $p \in L_x(M)$ is ***D*-adapted** if there is a splitting $E = E_1 \oplus E_2$ such that $p|E_1$ is an isomorphism onto D_x. If we *fix* the splitting $E = E_1 \oplus E_2$ we denote by $L(M, D)$ the associated collection of D-adapted linear frames. We observe that $L(M, D)$ is invariant under the action of the subgroup $GL(E; E_1)$ of $GL(E)$ consisting of those automorphisms which leave invariant E_1. A typical element of $GL(E; E_1)$ is of the form

$$\begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$$

where $a \in GL(E_1)$, $b \in GL(E_2)$, and $c \in L(E_2, E_1)$. If D' is a distribution complementary to D we say that $p \in L_x(M)$ is ***(D, D')-adapted*** if there exists a splitting $E = E_1 \oplus E_2$ and $p|E_1$ is an isomorphism onto D_x and $p|E_2$ is an isomorphism onto D'_x. As above, we fix the splitting $E = E_1 \oplus E_2$ and denote by $L(M, D, D')$ the set of (D, D')-adapted linear frames. Note that $L(M, D, D')$ is invariant under the action of $GL(E_1) \times GL(E_2)$ which we regard as a "diagonal" subgroup of $GL(E)$. In general terminology, $L(M, D)$ and $L(M, D, D')$ are ***subbundles*** of $L(M)$. We refer the reader to Section II.6 of [Kobayashi and Nomizu 1963] for a discussion of subbundles of principal bundles. We say that a linear connection on $L(M)$ ***restricts*** to $L(M, D)$ if its horizontal subspaces are tangent to $L(M, D)$. The following result is a basic one relating the notion of restriction of linear and affine connections.

4.6 Proposition: *Let D be a distribution on a manifold M with an affine connection ∇ and denote by $H(L(M))$ the associated linear connection. $H(L(M))$ restricts to $L(M, D)$ if and only if ∇ restricts to D.*

Proof: Let $E = E_1 \oplus E_2$ be the splitting associated with the construction of $L(M, D)$. Note that the function

$$\begin{aligned} \phi_Y \colon L(M) &\to E \\ p &\mapsto \theta(\mathrm{hlft}_p(Y)) = p^{-1}(Y(\pi(p))) \end{aligned}$$

when restricted to $L(M, D)$ takes its values in E_1 if $Y \in \mathscr{D}$. If $H(L(M))$ restricts to $L(M, D)$ it is then clear that $(\mathrm{hlft}\, X(\phi_Y))(p) \in E_1$ and so ∇ restricts to D by (2.1). Now suppose that ∇ restricts to D. It suffices to show that every standard horizontal vector field $\Sigma(e)$ is tangent to $L(M, D)$. So let $e \in E$, $p \in L(M, D)$, and denote by $\sigma \colon [0, T] \to L(M)$ the integral curve of $\Sigma(e)$ with $\sigma(0) = p$. Let $t \in]0, T]$ and let $v \in D_{c(t)}$. By parallel translating v *back* along c to $x = c(0)$ we obtain $u \in T_xM$. Since ∇ restricts to D, $u \in D_x$. Now let $e = p^{-1}u$. Since $p \in L(M, D)$ $e \in E_1$. Furthermore, since $\sigma(t)e = v$ and since $v \in D_{c(t)}$, $\sigma(t)$ must lie in $L(M, D)$. Thus $\Sigma(e)$ is tangent to $L(M, D)$ and this completes the proof. ■

We may now state our main result in this section.

4.7 Theorem: *Let M be a manifold modelled on a Banach space E, let D be a distribution on M with complement D', and let ∇ be an affine connection on M. Denote by $H(L(M))$ the linear connection associated with ∇ and let ω be the connection form. The following are equivalent:*

(i) $H(L(M))$ restricts to $L(M, D, D')$;

(ii) both D and D' are geodesically invariant;

(iii) there exists a splitting $E = E_1 \oplus E_2$ such that ω restricted to $L(M, D, D')$ takes its values in $\mathfrak{gl}(E_1) \oplus \mathfrak{gl}(E_2)$.

Proof: In the proof we let $E = E_1 \oplus E_2$ be a splitting such that if $p \in L_x(M)$ then $p|E_1$ is an isomorphism onto D_x and $p|E_2$ is an isomorphism onto D'_x.

(i) $\Longrightarrow$ (ii) If $H_p(L(M)) \subset T_p(L(M,D,D'))$ then $L(M,D,D')$ must be invariant under the flow of every horizontal vector field. Let $v \in D_x$ and let $p \in L_x(M,D,D')$. We then see that $e = p^{-1}v \in E_1$. Let $\sigma\colon [0,T] \to L(M)$ denote the integral curve of $\Sigma(e)$ with initial condition p and let $c = \pi \circ \sigma$ be the associated geodesic. Since $\Sigma(e)$ is horizontal, $\sigma(t) \in L(M,D,D')$ for $t \in]0,T]$. By definition of $\Sigma(e)$, $\dot{c}(t) = \sigma(t)e$. Therefore $\dot{c}(t) \in D_{c(t)}$ and so D is geodesically invariant. A similar argument shows that D' is geodesically invariant.

(ii) $\Longrightarrow$ (i) We shall show that every standard horizontal vector field leaves $L(M,D,D')$ invariant and this will suffice to show that $H_p(L(M)) \subset T_p(L(M,D,D'))$. Let $e \in E_1$ and let $\Sigma(e)$ be the associated standard horizontal vector field. Let $p \in L_x(M,D,D')$ and denote $v = pe \in D_x$. We let $c\colon [0,T] \to M$ be the geodesic with initial velocity v and let σ be the horizontal lift of c through p which is then the integral curve of $\Sigma(e)$ through p. Since D is geodesically invariant, $\dot{c}(t) \in D_{c(t)}$ for $t \in]0,T]$. Thus $\sigma(t)e \in D_{c(t)}$ since $\Sigma(e)$ is a standard horizontal vector field. Thus $\sigma(t) \in L_{c(t)}(M,D,D')$ and so $L(M,D,D')$ is invariant under the flow of $\Sigma(e)$. A similar argument for $e \in E_2$ using geodesic invariance of D' gives the desired result.

(ii) $\Longrightarrow$ (iii) Proposition 6.2 in Chapter II of [Kobayashi and Nomizu 1963] states that the connection form restricted to a subbundle takes its values in the Lie algebra of the reduced structure group. This is precisely what we have stated here in the special case we are considering.

(iii) $\Longrightarrow$ (ii) By Proposition 6.4 of Chapter II in [Kobayashi and Nomizu 1963] it suffices to determine a subspace of $\mathfrak{gl}(E)$ which is complementary to $\mathfrak{gl}(E_1) \oplus \mathfrak{gl}(E_2)$ and which is invariant under the adjoint action of $GL(E_1) \times GL(E_2)$. We claim that the subspace $\mathfrak{m}$ generated by endomorphisms of the form

$$C = \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$$

for $A \in L(E_2,E_1)$ and $B \in L(E_1,E_2)$ meets the criteria. Clearly $\mathfrak{gl}(E) = \mathfrak{gl}(E_1) \oplus \mathfrak{gl}(E_2) \oplus \mathfrak{m}$. Also, if

$$a \triangleq \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \in GL(E_1) \times GL(E_2)$$

then we compute

$$\mathrm{Ad}_a(C) = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix} \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & a_1 A a_2^{-1} \\ a_2 B A_1^{-1} & 0 \end{pmatrix} \in \mathfrak{m}. \qquad \blacksquare$$

4.8 Remarks: 1. In particular note that a sufficient condition for ∇ to restrict to a geodesically invariant distribution D is for there to exist a geodesically invariant complement to D. This condition is not necessary.

2. Note that if ∇ has zero torsion, then all the distributions in the theorem are integrable by Proposition 3.1.

3. Of course, the notion of being able to characterise restriction completely in terms of geodesic invariance is hopeless since geodesic invariance only depends upon the zero torsion part of the affine connection. However, stronger results than Theorem 4.7 should be possible.

4. In the case when ∇ is a Levi-Civita connection, the associated linear connection reduces to the "bundle of orthonormal frames" (see [Lang 1995, Section VII.3] for the infinite-dimensional version of this) and the structure group is $O(E)$ where we are supposing E to be equipped with a fixed inner product. Due to the orthogonal structure group, one can slightly sharpen Theorem 4.7 since $O(M, D, D^{\perp}) = O(M, D)$. •

5. Restricting affine connections to distributions

In this section we investigate various consequences of a natural restriction of a given affine connection to a distribution.

5.1. Motivation from mechanics. The basic construction we perform in this section comes from the dynamics of a class of mechanical systems with nonholonomic constraints. One has a Riemannian manifold (M, g) with ∇ the Levi-Civita affine connection. Additionally, one has a distribution D on M. The objective is to determine geodesics of ∇ which are subject to the constraint that their tangent vectors lie in D. The Lagrange-d'Alembert principle (see [Lanczos 1970])[4] states that the constrained solutions are those curves c which satisfy

$$\nabla_{\dot{c}(t)}\dot{c}(t) \in D^{\perp}_{c(t)}, \qquad \dot{c}(t) \in D_{c(t)}$$

where $D^{\perp}$ is the g-orthogonal complement to D. Equivalently, there exists a section λ of $D^{\perp}$ along c such that

$$\nabla_{\dot{c}(t)}\dot{c}(t) = \lambda(t) \tag{5.1a}$$

$$P'(\dot{c}(t)) = 0 \tag{5.1b}$$

where $P' \colon TM \to TM$ is the orthogonal projection onto $D^{\perp}$. One may differentiate the constraint equation (5.1b) to obtain

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

Applying P' to the equations of motion (5.1a) we obtain

$$P'(\nabla_{\dot{c}(t)}\dot{c}(t)) = \lambda(t)$$

as $\lambda(t) \in D^{\perp}_{c(t)}$. Combining the two equations to eliminate λ one sees that c must be a geodesic of the affine connection $\overline{\nabla}$ which is defined by

$$\overline{\nabla}_X Y = \nabla_X Y + (\nabla_X P')(Y).$$

Conversely, if $c \colon [a, b] \to M$ is a geodesic for $\overline{\nabla}$ with $\dot{c}(a) \in D_{c(a)}$ then c satisfies (5.1). This is a version of the approach taken by Synge [1928]. Other authors have subsequently taken

[4] There are actually several ways to write equations for the systems we are considering. The method we choose has the property of agreeing with Newton's equations in instances where both methods apply. We refer the reader to [Lewis and Murray 1995] for a discussion and critique of another method of deriving the constrained equations.

up various incarnations of this method. We mention in particular [Cattaneo-Gasparini 1963], [Cattaneo 1963], [Vershik 1984], and [Bloch and Crouch 1995].[5]

The remainder of this section will be devoted to generalising this construction and understanding the properties of the new affine connection $\overline{\nabla}$.

5.2. Generalising the second fundamental form. The second fundamental form for submanifolds arises when one restricts a given affine connection. We shall specify a way of restricting an affine connection to a distribution and in doing so, compute an associated second fundamental form.

We fix a distribution D on M. Recall that when one computes the second fundamental form in Riemannian submanifold geometry, one uses the Riemannian metric to specify directions normal to the submanifold. We refer the reader to [Nomizu and Sasaki 1994] for a discussion of the second fundamental form for submanifolds in the situation when the affine connection is not a Levi-Civita connection. In such cases, the construction depends on the choice of a normal space. Thus we *fix* a complement D' to D. Let us denote by $P\colon TM \to TM$ the projection onto D and by $P'\colon TM \to TM$ the projection onto D'. We will also think of P and P' as $(1,1)$ tensor fields on M. We shall assume that M comes equipped with an affine connection ∇ which need not be Levi-Civita. We then define a new affine connection $\overline{\nabla}$ on M by

$$\overline{\nabla}_X Y = \nabla_X Y + (\nabla_X P')(Y).$$

That this is indeed an affine connection follows since $(X, Y) \mapsto (\nabla_X P')(Y)$ is $C^\infty(M)$-bilinear. We have some important properties of $\overline{\nabla}$.

5.1 Proposition: *(i)* $\overline{\nabla}_X Y \in \mathscr{D}$ *for every* $X \in \mathscr{T}(M)$ *and* $Y \in \mathscr{D}$,

(ii) $(\nabla_X P')(Y) \in \mathscr{D}'$ *for* $X \in \mathscr{T}(M)$ *and* $Y \in \mathscr{D}$, *and*

(iii) $(\nabla_X P')(Y) \in \mathscr{D}$ *for* $X \in \mathscr{T}(M)$ *and* $Y \in \mathscr{D}'$.

Proof: (i) Let X and Y be vector fields on M. Then

$$P'(\overline{\nabla}_X Y) = P'(\nabla_X Y) + P'(\nabla_X P')(Y). \tag{5.2}$$

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

$$\begin{aligned} & P'(Y) = 0 \\ \Longrightarrow\quad & (\nabla_X P')(Y) + P'(\nabla_X Y) = 0 \qquad (5.3) \\ \Longrightarrow\quad & P'(\nabla_X P')(Y) + P'(\nabla_X Y) = 0 \qquad (5.4) \end{aligned}$$

since $P' \circ P' = P'$. Substituting (5.4) into (5.2) we see that $P'(\overline{\nabla}_X Y) = 0$ for $X \in \mathscr{T}(M)$ and $Y \in \mathscr{D}$. Therefore, $\overline{\nabla}_X Y \in \mathscr{D}$.

(ii) Let $Y \in \mathscr{D}$. From (5.3) we have

$$\begin{aligned} & (\nabla_X P')(Y) + P'(\nabla_X Y) = 0 \\ \Longrightarrow\quad & P(\nabla_X P')(Y) = 0 \end{aligned}$$

[5]The author is indebted to M. Favretti for pointing out an unpublished paper by S. Benenti which contains the first two references.

since $P \circ P' = 0$. Thus $(\nabla_X P')(Y) \in \mathscr{D}'$ for $Y \in \mathscr{D}$.

(iii) Let $Y \in \mathscr{D}'$. Then

$$\begin{aligned} & P'(Y) = Y \\ \implies \quad & (\nabla_X P')(Y) + P'(\nabla_X Y) = \nabla_X Y \\ \implies \quad & P'(\nabla_X P')(Y) + P'(\nabla_X Y) = P'(\nabla_X Y) \\ \implies \quad & P'(\nabla_X P')(Y) = 0. \end{aligned}$$

Here we have used the fact that $P' \circ P' = P'$. This shows that $(\nabla_X P')(Y) \in \mathscr{D}$ if $Y \in \mathscr{D}'$. ■

Note that (i) implies that $\overline{\nabla}$ restricts to D. In particular we see that D is geodesically invariant with respect to $\overline{\nabla}$ by Theorem 4.4(iii). We shall see below that it is properties (i) and (ii) which are the most interesting.

The above result has an important consequence. For $Y \in \mathscr{D}$ note that the expression $\nabla_X Y = \overline{\nabla}_X Y - (\nabla_X P)(Y)$ is simply the decomposition of $\nabla_X Y$ into its D and D' components. The D component may be thought of as the restriction of ∇ to D. Following what one does in the submanifold case, the remainder (i.e., the D' component), must be the second fundamental form. We make the following definition.

5.2 Definition: Let D be a distribution on a manifold M with affine connection ∇ and let D' be a distribution complementary to D. The section S of $L^2(D, D')$ defined by

$$S(u, v) = -(\nabla_u P')(v)$$

is the ***second fundamental form*** for (∇, D, D'). •

The definition *does* depend upon the choice of a complement D' to D so this should be included in the definition of the affine second fundamental form. In the sequel we will find it convenient to have some notation for the term $(\nabla_X P')(Y)$ when X and Y are not sections of D. Thus we define the $(1, 2)$ tensor

$$Q\colon (X, Y) \mapsto (\nabla_X P')(Y). \tag{5.5}$$

The following result generalises a classical result from submanifold geometry.

5.3 Proposition: *D is geodesically invariant under ∇ if and only if S is skew-symmetric.*

Proof: First suppose that D is geodesically invariant and let $X, Y \in \mathscr{D}$. We may write

$$\nabla_X Y + \nabla_Y X = \overline{\nabla}_X Y + S(X, Y) + \overline{\nabla}_Y X + S(Y, X).$$

By Theorem 4.4(ii) $\nabla_X Y + \nabla_Y X \in \mathscr{D}$. By Proposition 5.1(i) and (ii), respectively,

$$\overline{\nabla}_X Y, \overline{\nabla}_Y X \in \mathscr{D}, \qquad S(X, Y), S(Y, X) \in \mathscr{D}'.$$

Therefore $S(X, Y) + S(Y, X) = 0$ for every $X, Y \in \mathscr{D}$. Now suppose that S is skew-symmetric. For $X \in \mathscr{D}$ we then have

$$\nabla_X X = \overline{\nabla}_X X + S(X, X) = \overline{\nabla}_X X.$$

By Proposition 5.1(i) we have $\nabla_X X \in \mathscr{D}$. This implies that D is geodesically invariant by Theorem 4.4(iii). ■

5.4 Remarks: 1. In particular note that if $S = 0$ then D is geodesically invariant. The converse of this is not true in general. One needs to add the hypotheses that D be integrable and that ∇ be torsion free. We shall see below how this works in the Levi-Civita case (cf. Corollary 5.6).

2. The notion of D being geodesically invariant has nothing to do with the choice of complement D'. However, the second fundamental form which we have defined *does* depend on D'. The above result may then be interpreted as saying that D is geodesically invariant if and only if S is skew-symmetric *for any choice of complement* D'. •

Next we show that $\overline{\nabla}$ does in fact generalise what one normally does in the Levi-Civita case when restricting the connection to a submanifold. Because our definitions only make sense for distributions, we extend the classical construction by working with regular integrable distributions rather than submanifolds. In this case we choose D' to be $D^{\perp}$, the orthogonal complement of D. Thus S takes its values in $D^{\perp}$.

5.5 Proposition: *Suppose that D is integrable and let Λ be a maximal integral manifold which we suppose to be an embedded submanifold of M. Denote by g_Λ the Riemannian metric on Λ induced by g and denote by $\overset{\Lambda}{\nabla}$ the associated Levi-Civita connection on Λ. Then $\overline{\nabla}_X Y = \overset{\Lambda}{\nabla}_X Y$ for $X, Y \in \mathscr{T}(\Lambda)$.*

Proof: In the proof, when we write vector fields on Λ, we shall consider them to be extended to M when necessary. One may verify that the results do not depend on the choice of extension.

We must show that $\overline{\nabla}$ restricted to $\mathscr{T}(\Lambda)$ is torsion free and preserves the metric g_Λ. We first show that $\overline{\nabla}$ is torsion free. Let $X, Y \in \mathscr{T}(\Lambda)$. We have

$$\overline{\nabla}_X Y - \overline{\nabla}_Y X = \nabla_X Y - \nabla_Y X - S(X,Y) + S(Y,X).$$

By Proposition 5.1(ii) we see that $S(X,Y)$ and $S(Y,X)$ are normal to Λ, and by Proposition 5.1(i) we see that $\overline{\nabla}_X Y, \overline{\nabla}_Y X \in \mathscr{T}(\Lambda)$. Also, since ∇ is torsion free and since D is integrable,

$$\nabla_X Y - \nabla_Y X = [X,Y] \in \mathscr{T}(\Lambda).$$

Thus we obtain

$$\overline{\nabla}_X Y - \overline{\nabla}_Y X = [X,Y], \qquad S(X,Y) = S(Y,X)$$

which shows in particular that $\overline{\nabla}$ is torsion free.

Now we show that $\overline{\nabla}$ preserves g_Λ. Let $X, Y, Z \in \mathscr{T}(\Lambda)$. Since ∇ is a Levi-Civita connection,

$$\mathscr{L}_X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

Now we have

$$g(\nabla_X Y, Z) = g(\overline{\nabla}_X Y, Z) + g(S(X,Y), Z) = g(\overline{\nabla}_X Y, Z)$$

since $S(X,Y)$ is normal to Λ. Similarly we may show that

$$g(Y, \nabla_X Z) = g(Y, \overline{\nabla}_X Z).$$

Therefore

$$\mathscr{L}_X(g(Y,Z)) = g(\overline{\nabla}_X Y, Z) + g(Y, \overline{\nabla}_X Z)$$

which, when restricted to Λ, shows that $\overline{\nabla}$ preserves g_Λ. ■

In the proof of the proposition we have also derived the following result.

5.6 Corollary: *With the hypotheses of the Proposition 5.5 we have $S(X,Y) = S(Y,X)$ for $X, Y \in \mathscr{T}(\Lambda)$.*

In Riemannian geometry this is the classical statement that the second fundamental form is symmetric.

5.3. Affine transformations for $\overline{\nabla}$. We now investigate certain affine transformations of $\overline{\nabla}$. We are interested in determining when an affine transformation for ∇ is a D-affine transformation for $\overline{\nabla}$.

5.7 Proposition: *Let ϕ be an affine transformation for ∇ which is compatible with (D, D') and let X be an infinitesimal affine transformation for ∇ which is compatible with (D, D'). Then*

(i) ϕ is a D-affine transformation for $\overline{\nabla}$, and

(ii) X is a D-infinitesimal affine transformation for $\overline{\nabla}$.

Proof: (i) Let $Y \in \mathscr{D}$. We have

$$\begin{aligned} \phi^*(\overline{\nabla}_X Y) &= \phi^*\left(\nabla_X Y + Q(X,Y)\right) \\ &= \nabla_{\phi^* X}\phi^* Y + \phi^*(Q(X,Y)) \end{aligned} \tag{5.6}$$

since ϕ is an affine transformation for ∇. By Proposition 5.1(i) and since ϕ is compatible with D we have $\phi^*(\overline{\nabla}_X Y) \in \mathscr{D}$. By Proposition 5.1(ii) and since ϕ is compatible with D' we have $\phi^*(Q(X,Y)) \in \mathscr{D}'$. Therefore, (5.6) is simply the decomposition of $\nabla_{\phi^* X}\phi^* Y$ into its D and D' components. Therefore, by Proposition 5.1 we must have

$$\phi^*(\overline{\nabla}_X Y) = \overline{\nabla}_{\phi^* X}\phi^* Y \quad \text{and} \quad \phi^*(Q(X,Y)) = Q(\phi^* X, \phi^* Y).$$

In particular, ϕ is a D-affine transformation by Proposition 3.8(iii).

(ii) Let $\bar{B}_X$ denote the $(1,2)$ tensor field associated with $\overline{\nabla}$ (cf. equation (2.3)). We compute

$$\begin{aligned} \bar{B}_X(Y,Z) = [X, \nabla_Y Z] + [X, Q(Y,Z)] - \nabla_Y[X,Z] - \\ Q(Y,[X,Z]) - \nabla_{[X,Y]}Z - Q([X,Y],Z). \end{aligned}$$

If X is an infinitesimal affine transformation for ∇ then

$$[X, \nabla_Y Z] - \nabla_Y[X,Z] - \nabla_{[X,Y]}Z = 0.$$

Therefore

$$\bar{B}_X(Y,Z) = [X, Q(Y,Z)] - Q(Y,[X,Z]) - Q([X,Y],Z).$$

We now suppose that $Y, Z \in \mathscr{D}$. By Proposition 5.1(ii), $Q(Y,Z) \in \mathscr{D}'$ and since X is compatible with D', $[X, Q(Y,Z)] \in \mathscr{D}'$. Similar arguments using Proposition 5.1(ii) and the fact that X is compatible with D show that

$$Q(Y,[X,Z]), Q([X,Y],Z) \in \mathscr{D}'.$$

Therefore $\bar{B}_X(Y,Z) \in \mathscr{D}'$. However, by Lemma 3.6 we know that $\bar{B}_X(Y,Z) \in \mathscr{D}$. Therefore, $\bar{B}_X(Y,Z) = 0$ and so X is a D-infinitesimal affine transformation by Proposition 3.8(vi). ■

We denote by $\mathit{Aff}_0(\overline{\nabla})$ the set of D-affine transformations of $\overline{\nabla}$ obtained as in Proposition 5.7(i) and by $\mathfrak{aff}_0(\overline{\nabla})$ the set of D-infinitesimal affine transformations of $\overline{\nabla}$ obtained as in Proposition 5.7(ii). The following result puts $\mathit{Aff}_0(\overline{\nabla})$ in its place with respect to $\mathit{Aff}(\nabla)$, $\mathit{Aff}(\overline{\nabla})$, and $\mathit{Aff}(\overline{\nabla}|D)$ (and similarly for $\mathfrak{aff}_0(\overline{\nabla})$).

5.8 Proposition: *(i) $\mathit{Aff}_0(\overline{\nabla})$ is a subgroup of $\mathit{Aff}(\nabla)$, $\mathit{Aff}(\overline{\nabla})$, and $\mathit{Aff}(\overline{\nabla}|D)$, and*
(ii) $\mathfrak{aff}_0(\overline{\nabla})$ is a Lie subalgebra of $\mathfrak{aff}(\nabla)$, $\mathfrak{aff}(\overline{\nabla})$, and $\mathfrak{aff}(\overline{\nabla}|D)$.

Proof: That $\mathit{Aff}_0(\overline{\nabla})$ is a subgroup of $\mathit{Diff}(M)$ (resp. $\mathfrak{aff}_0(\overline{\nabla})$ is a Lie subalgebra of $\mathscr{T}(M)$) follows from the fact that diffeomorphisms compatible with D are a subgroup of $\mathit{Diff}(M)$ (resp. a Lie subalgebra of $\mathscr{T}(M)$) as was shown in the proof of Proposition 3.9. The proposition now follows since the inclusions as sets are obvious. ■

5.4. Conservation laws for restricted Levi-Civita geodesics. In this section we examine the results of Section 5.3 in the case when ∇ is the Levi-Civita connection associated with a Riemannian metric g on M. We first look at conservation laws for the restricted system. We are able to give conditions under which a conserved quantity for the unrestricted system will persist for the restricted system. Then we describe the momentum equation of Bloch, Krishnaprasad, Marsden, and Murray [1996].

Unless otherwise stated, in this section ∇ will be the Levi-Civita connection associated with a Riemannian metric g.

Conserved quantities. By virtue of Lemma 3.10, the conclusions of Proposition 5.7 will always apply in this case. That is to say, every affine (resp. infinitesimal affine) transformation of ∇ which is compatible with D is a D-affine (resp. D-infinitesimal affine) transformation of $\overline{\nabla}$. However, in the case when M has a Riemannian metric, infinitesimal affine transformations of the Levi-Civita connection lead to conserved quantities for the geodesic flow. The following result is a special case of what is known as Noether's Theorem in the mechanics literature [Abraham and Marsden 1978, Corollary 4.2.14].

5.9 Proposition: *Let ∇ be the Levi-Civita connection associated with a Riemannian metric g on M and let X be a vector field on M. Then X is an infinitesimal isometry associated with g if and only if the function*

$$\boldsymbol{J}_X(v_x) = g(X(x), v_x)$$

*on TM is constant along the integral curves of the geodesic spray. We shall call $\boldsymbol{J}_X$ the **momentum** associated with X.*

Proof: Let c be a geodesic of ∇. We compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{J}_X(\dot{c}(t)) &= (\nabla_{\dot{c}(t)}g)(X(c(t)), \dot{c}(t)) + g(\nabla_{\dot{c}(t)}X(c(t)), \dot{c}(t)) + g(X(c(t)), \nabla_{\dot{c}(t)}\dot{c}(t)) \\ &= g(\nabla_{\dot{c}(t)}X(c(t)), \dot{c}(t)).\end{aligned}$$

Therefore, $\boldsymbol{J}_X$ is constant along every geodesic if and only if

$$g(\nabla_{\dot{c}(t)}X(c(t)), \dot{c}(t)) = 0$$

for every geodesic c of ∇. The proposition now follows from Lemma 2.1(iii). ■

Note that we can define the momentum associated with *any* vector field Y by $\boldsymbol{J}_Y(v_x) = g(Y(x), v_x)$. However, it is only the momenta associated with Killing vector fields which lead to conserved quantities.

It is possible to state conditions under which an infinitesimal isometry for g gives rise to a conserved quantity along the geodesics of $\overline{\nabla}$ whose initial velocities lie in D. We first state two computational lemmas.

5.10 Lemma: *Let X, Y, Z be vector fields on M. Then*

$$(\overline{\nabla}_X g)(Y, Z) = -g(Q(X, Y), Z) - g(Y, Q(X, Z)).$$

Proof: We compute

$$\mathscr{L}_X(g(Y, Z)) = (\nabla_X g)(Y, Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \tag{5.7}$$

Also

$$\begin{aligned}\mathscr{L}_X(g(Y, Z)) &= (\overline{\nabla}_X g)(Y, Z) + g(\overline{\nabla}_X Y, Z) + g(Y, \overline{\nabla}_X Z)\\ &= (\overline{\nabla}_X g)(Y, Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z) + g(Q(X, Y), Z) + g(Y, Q(X, Z)).\end{aligned} \tag{5.8}$$

Subtracting (5.7) from (5.8) and using the fact that $\nabla_X g = 0$ we obtain the desired result. ■

5.11 Lemma: *Let X be an infinitesimal affine transformation for ∇. Then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{J}_X(\dot{c}(t)) = g(X(c(t)), S(\dot{c}(t), \dot{c}(t)))$$

for every geodesic c of $\overline{\nabla}$ whose initial velocity lies in D.

Proof: Let c be a geodesic of $\overline{\nabla}$ whose initial velocity lies in D. We compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{J}_X(\dot{c}(t)) &= (\overline{\nabla}_{\dot{c}(t)} g)(X(c(t)), \dot{c}(t)) + g(\overline{\nabla}_{\dot{c}(t)} X(c(t)), \dot{c}(t)) + g(X(c(t)), \overline{\nabla}_{\dot{c}(t)}\dot{c}(t))\\ &= (\overline{\nabla}_{\dot{c}(t)} g)(X(c(t)), \dot{c}(t)) + g(\nabla_{\dot{c}(t)} X(c(t)), \dot{c}(t)) + g(Q(\dot{c}(t), X(c(t))), \dot{c}(t))\\ &= (\overline{\nabla}_{\dot{c}(t)} g)(X(c(t)), \dot{c}(t)) + g(Q(\dot{c}(t), X(c(t))), \dot{c}(t))\\ &= g(Q(\dot{c}(t), X(c(t))), \dot{c}(t)) - g(Q(\dot{c}(t), X(c(t))), \dot{c}(t)) - g(X(c(t)), Q(\dot{c}(t), \dot{c}(t)))\\ &= -g(X(c(t)), Q(\dot{c}(t), \dot{c}(t))).\end{aligned}$$

Here we have used the fact that X is an infinitesimal affine transformation of ∇, the fact that c is a geodesic of $\overline{\nabla}$, Proposition 5.1(ii), and Lemma 5.10. The lemma follows since $Q|D = -S$. ■

As a consequence of these computations, we have the following result.

5.12 Proposition: *Let X be an infinitesimal affine transformation for ∇ and suppose that $X \in \mathscr{D}$. Then $\boldsymbol{J}_X$ is constant along geodesics of $\overline{\nabla}$ whose initial velocities lie in D.*

Proof: Apply Proposition 5.1(ii) and Lemma 5.11. ■

Notice that in order for an infinitesimal isometry X to lead to a conservation law for the restricted system, it is not necessary that X be compatible with the distribution. We also mention that it is still possible that $\boldsymbol{J}_X$ be conserved even if $X \notin \mathscr{D}$. We shall see both of these phenomenon exhibited in the example we consider in Section 5.6.

The momentum equation. Now we present an interpretation of the momentum equation of Bloch, Krishnaprasad, Marsden, and Murray [1996] in the context of our constructions. The momentum equation may be thought of as describing the evolution of the momenta of the unrestricted system when restricted to the distribution D. The following result makes this precise.

5.13 Proposition: *Let $X_1, \ldots, X_m$ be Killing vector fields for g. Suppose that there exist functions $\gamma^1, \ldots, \gamma^m$ on M so that the vector field*

$$Y = \gamma^1 X_1 + \cdots + \gamma^m X_m$$

is a section of D. Then the momentum associated with Y satisfies the evolution equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{J}_Y(\dot{c}(t)) = g(\dot{\gamma}^a X_a(c(t)), \dot{c}(t)) \tag{5.9}$$

along geodesics c of $\overline{\nabla}$.

Proof: We compute

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{J}_Y(c(t)) &= (\overline{\nabla}_{\dot{c}(t)} g)(Y(c(t)), \dot{c}(t)) + g(\overline{\nabla}_{\dot{c}(t)} Y(c(t)), \dot{c}(t)) + g(Y(c(t)), \overline{\nabla}_{\dot{c}(t)}\dot{c}(t)) \\
&= (\overline{\nabla}_{\dot{c}(t)} g)(Y(c(t)), \dot{c}(t)) + \gamma^a(c(t)) g(\nabla_{\dot{c}(t)} X_a(c(t)), \dot{c}(t)) + \\
&\quad g(Q(\dot{c}(t), Y(c(t))), \dot{c}(t)) + \dot{\gamma}^a(c(t)) g(X_a(c(t)), \dot{c}(t)) \\
&= -g(Q(\dot{c}(t), Y(c(t))), \dot{c}(t)) - g(Y(c(t)), Q(\dot{c}(t), \dot{c}(t))) + \\
&\quad g(Q(\dot{c}(t), Y(c(t))), \dot{c}(t)) + \dot{\gamma}^a(c(t)) g(X_a(c(t)), \dot{c}(t)) \\
&= -g(Y(c(t)), Q(\dot{c}(t), \dot{c}(t))) + \dot{\gamma}^a(c(t)) g(X_a(c(t)), \dot{c}(t)) \\
&= \dot{\gamma}^a(c(t)) g(X_a(c(t)), \dot{c}(t)).
\end{aligned}$$

Here we have used the fact that X_a, $a = 1, \ldots, m$ are Killing vector fields, c is a geodesic of $\overline{\nabla}$, $Y \in \mathscr{D}$, and Proposition 5.1(ii) and Lemma 5.10. ■

It is equation (5.9) which is the momentum equation described by Bloch, Krishnaprasad, Marsden, and Murray. It is interesting to note that, although the properties of $\overline{\nabla}$ were used in the derivation of the momentum equation, $\overline{\nabla}$ does not appear in its final form.

5.5. Extending connections from D to TM. In this section we take the point of view of treating $\overline{\nabla}|D$ as the given object and characterise those affine connections on all of TM which agree with $\overline{\nabla}$ when restricted to D. This point of view is brought up by Vershik [1984] and Bloch and Crouch [1995], although in each case the context is somewhat different than ours.

5.14 Proposition: *Let D be a distribution on a manifold M and suppose that D' is a complement to D. Let ∇ be an affine connection on M and suppose that another affine connection $\widetilde{\nabla}$ has the properties*

(i) $\widetilde{\nabla}_X Y \in \mathscr{D}$ for every $Y \in \mathscr{D}$, and

(ii) $\widetilde{\nabla}_X Y - \nabla_X Y \in \mathscr{D}'$ for every $Y \in \mathscr{D}$.

Then $\widetilde{\nabla}_X Y = \nabla_X Y + (\nabla_X P')(Y) + S(X,Y)$ for some $(1,2)$ tensor field S such that $P'(S(X,Y)) = 0$ for $Y \in \mathscr{D}$. Conversely, if $\widetilde{\nabla}$ is of this form, then it satisfies (i) and (ii).

Proof: We may write *any* affine connection on M as

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

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

$$\begin{aligned} & P'(Y) = 0 \\ \implies \quad & (\nabla_X P')(Y) + P'(\nabla_X Y) = 0 \\ \implies \quad & P'(\nabla_X Y) = -(\nabla_X P')(Y). \end{aligned} \tag{5.10}$$

We also have

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

Using (i), (ii), and (5.10) we have

$$P'(\nabla_X Y) + B(X,Y) = 0 \quad \implies \quad B(X,Y) = (\nabla_X P')(Y).$$

Thus $\widetilde{\nabla}_X Y = \nabla_X Y + (\nabla_X P')(Y) + S(X,Y)$ for some S such that $P'(S(X,Y)) = 0$ for $Y \in \mathscr{D}$. The final assertion follows easily by Proposition 5.1. ■

5.15 Remark: It is important to note that those things we have said about $\overline{\nabla}$ in this section which only depend on its restriction to D will be equally true for *any* of the affine connections described by Proposition 5.14. This is true in particular for our discussion of transformations in Section 5.3 and conservation laws in Section 5.4. •

To conclude this section, we provide an affine connection which restricts to D and whose restriction to D is the same as $\overline{\nabla}$. The connection we construct, however, has the property of being easier to compute in some examples, so it is worth recording. If A is an arbitrary invertible $(1,1)$ tensor field on M we define an affine connection on M by

$$\overset{A}{\nabla}_X Y = \nabla_X Y + A^{-1}(\nabla_X(AP'))(Y). \tag{5.11}$$

We claim that $\overset{A}{\nabla}$ satisfies properties (i) and (ii) of Proposition 5.14. We compute

$$A^{-1}(\nabla_X(AP'))(Y) = A^{-1}(\nabla_X A)(P'(Y)) + (\nabla_X P')(Y).$$

Observing that $P'(A^{-1}(\nabla_X A)(P'(Y))) = 0$ for $Y \in \mathscr{D}$ we see that $\overset{A}{\nabla}$ does indeed satisfy properties (i) and (ii) of Proposition 5.14 by applying the final assertion of that proposition.

In examples one can use A to simplify the term AP' *before* it gets differentiated and this is often helpful. We shall see this in the example of Section 5.6. Note that if $A = \mathrm{id}_{TM}$ then $\overset{A}{\nabla} = \overline{\nabla}$.

5.6. A simple example. Here we take $M = \mathbb{R}^3$ and consider the Riemannian metric

$$g = \mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z.$$

We let ∇ be the associated Levi-Civita affine connection. The distribution D we consider is the span of the two vector fields

$$\frac{\partial}{\partial y}, \quad \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$$

on $\mathbb{R}^3$. This system is sometimes called the "Heisenberg system" since the brackets of the vector fields lying in D obey commutator relations which are reminiscent of the Lie algebra of the Heisenberg group.

The restricted connection. The orthogonal complement of D is easily seen to be generated by the vector field

$$y\frac{\partial}{\partial x} - \frac{\partial}{\partial z}.$$

We may easily compute P' to be

$$P'(x, y, z) \cdot (v_x, v_y, v_z) = \frac{1}{1+y^2}\left(y^2 v_x - y v_z, 0, -y v_x + v_z\right).$$

We shall use the extended connection $\overset{A}{\nabla}$ defined by (5.11) and choose $A = (1 + y^2)\,\mathrm{id}_{TM}$. We compute the non-zero connection coefficients of $\overset{A}{\nabla}$ to be

$$\Gamma^x_{xy} = \frac{2y}{1+y^2}, \qquad \Gamma^x_{zy} = -\frac{1}{1+y^2}, \qquad \Gamma^z_{xy} = -\frac{1}{1+y^2}.$$

Remember that the connection is not torsion free and so does not have the indicial symmetries of a Levi-Civita connection. The equations for the geodesics of this affine connection are

$$\begin{aligned}
&\ddot{x} + \frac{1}{1+y^2}\left(2y\dot{x}\dot{y} - \dot{y}\dot{z}\right) = 0\\
&\ddot{y} = 0\\
&\ddot{z} - \frac{1}{1+y^2}\dot{x}\dot{y} = 0.
\end{aligned}$$

These equations should be restricted to D since those are the only interesting initial conditions. The restricted velocities satisfy the equation

$$\dot{z} = y\dot{x}.$$

We thus compute the equations of motion restricted to D to be

$$\begin{aligned}
&\ddot{x} + \frac{y}{1+y^2}\dot{x}\dot{y} = 0\\
&\ddot{y} = 0\\
&\ddot{z} - \frac{1}{1+y^2}\dot{x}\dot{y} = 0.
\end{aligned}$$

***D*-affine transformations.** Now we turn to investigating the affine transformations of $\overset{A}{\nabla}$. Following Proposition 5.7 we seek affine transformations of ∇ which are compatible with D. It is easiest to look at the infinitesimal case. The Lie algebra of infinitesimal isometries of g is isomorphic to $\mathfrak{se}(3)$. The vector fields

$$X_1 = \frac{\partial}{\partial x}, \qquad X_2 = \frac{\partial}{\partial y}, \qquad X_3 = \frac{\partial}{\partial z},$$
$$X_4 = -z\frac{\partial}{\partial y} + y\frac{\partial}{\partial z}, \qquad X_5 = z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}, \qquad X_6 = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$$

form a basis for the Lie algebra of infinitesimal isometries of g. Of these vector fields, we may verify that only X_1 and X_3 are compatible with D. Therefore, both X_1 and X_3 are D-infinitesimal affine isometries of $\overset{A}{\nabla}$ by Proposition 5.7. The group generated by these transformations is isomorphic to $\mathbb{R}^2$ and the action on $\mathbb{R}^3$ is by $((a,b),(x,y,z)) \mapsto (x+a, y, z+b)$. Thus, in this example, $\mathit{Aff}_0(\overset{A}{\nabla})$ is a finite-dimensional Lie group and $\mathfrak{aff}_0(\overset{A}{\nabla})$ is its Lie algebra.

Evolution of momenta. We compute the conserved quantities of the unrestricted dynamics to be

$$\boldsymbol{J}_{X_1} = v_x, \qquad \boldsymbol{J}_{X_2} = v_y, \qquad \boldsymbol{J}_{X_3} = v_z,$$
$$\boldsymbol{J}_{X_4} = -zv_y + yv_z, \qquad \boldsymbol{J}_{X_5} = zv_x - xv_z, \qquad \boldsymbol{J}_{X_6} = -yv_x + xv_y.$$

Using Lemma 5.11 we derive that along the geodesics of $\overset{A}{\nabla}$ we have

$$\dot{\boldsymbol{J}}_{X_1} = \frac{\dot y\,(\dot z - 2y\dot x)}{1+y^2}, \qquad \dot{\boldsymbol{J}}_{X_2} = 0, \qquad \dot{\boldsymbol{J}}_{X_3} = \frac{\dot x\dot y}{1+y^2},$$
$$\dot{\boldsymbol{J}}_{X_4} = \frac{y\dot x\dot y}{1+y^2}, \qquad \dot{\boldsymbol{J}}_{X_5} = \frac{\dot y\,(z\dot z - (x+2yz)\dot x)}{1+y^2}, \qquad \dot{\boldsymbol{J}}_{X_6} = \frac{y\dot y\,(2y\dot x - \dot z)}{1+y^2}.$$

We may simplify these evolution equations by noting that we are only interested in those solutions lying in D. Substituting the restricted velocities into the equations for the evolution of the momenta we obtain

$$\dot{\boldsymbol{J}}_{X_1} = -\frac{y\dot x\dot y}{1+y^2}, \qquad \dot{\boldsymbol{J}}_{X_2} = 0, \qquad \dot{\boldsymbol{J}}_{X_3} = \frac{\dot x\dot y}{1+y^2},$$
$$\dot{\boldsymbol{J}}_{X_4} = \frac{y\dot x\dot y}{1+y^2}, \qquad \dot{\boldsymbol{J}}_{X_5} = -\frac{\dot x\dot y\,(x+yz)}{1+y^2}, \qquad \dot{\boldsymbol{J}}_{X_6} = \frac{y^2\dot x\dot y}{1+y^2}.$$

Looking at these relations, we see two conservation laws. First of all, since $X_2 \in \mathscr{D}$, the fact that $\boldsymbol{J}_{X_2}$ is conserved comes to us from Proposition 5.12. We also have the other obvious conservation law

$$\dot{\boldsymbol{J}}_{X_1+X_4} = 0.$$

The vector field $X_1 + X_4$ is not a section of D, so this conservation law is not a consequence of Proposition 5.12.

The momentum equation. Our definition of the momentum equation is fairly versatile, so we can write down many momentum equations. However, to illustrate the theory, we choose the same momentum equation for this example as was chosen by Bloch, Krishnaprasad, Marsden, and Murray [1996]. Thus we define a section of D by

$$Y = X_1 + yX_3.$$

This section of D is distinguished by being a linear combination of infinitesimal affine isometries which are compatible with D. Using Proposition 5.13 we derive

$$\dot{\boldsymbol{J}}_Y = g(\dot{y}X_3, \dot{x}X_1 + \dot{y}X_2 + \dot{z}X_3) = \dot{y}\dot{z}.$$

References

Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, 2nd edition, Addison Wesley: Reading, MA, ISBN: 978-0-8218-4438-0.

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.

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.

Bloch, A. M., Krishnaprasad, P. S., Marsden, J. E., and Murray, R. M. [1996] *Nonholonomic mechanical systems with symmetry*, Archive for Rational Mechanics and Analysis, **136**(1), pages 21–99, ISSN: 0003-9527, DOI: 10.1007/BF02199365.

Cattaneo, C. [1963] *Sulla struttura locale delle equazioni dinamiche di un sistema anholomo*, Atti della Accademia Nazionale dei Lincei. Rendiconti. Classe di Scienze Fisiche, Matematiche e Naturali. Serie VIII, **34**, pages 396–403, ISSN: 0392-7881.

Cattaneo-Gasparini, I. [1963] *Dérivée covariante "liée" dans une V_{n-1} riemannienne à structure presque produit*, Comptes Rendus Mathématique. Académie des Sciences, **256**, pages 2089–2091, ISSN: 1631-073X.

Crouch, P. E. [1981] *Geometric structures in systems theory*, Institution of Electrical Engineers. Proceedings. D. Control Theory and Applications, **128**(5), pages 242–252, ISSN: 1350-2379, DOI: 10.1049/ip-d:19810051.

Kobayashi, S. and Nomizu, K. [1963] *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 1996].

— [1996] *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 1963].

Lanczos, C. [1949] *The Variational Principles of Mechanics*, University of Toronto Press: Toronto, ON, Reprint: [Lanczos 1970].

— [1970] *The Variational Principles of Mechanics*, 4th edition, Dover Publications, Inc.: New York, NY, ISBN: 978-0-486-65067-8, Original: [Lanczos 1949].

Lang, S. [1995] *Differential and Riemannian Manifolds*, number 160 in Graduate Texts in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-94338-1.

Lewis, A. D. and Murray, R. M. [1995] *Variational principles for constrained systems: Theory and experiment*, International Journal of Non-Linear Mechanics, **30**(6), pages 793–815, ISSN: 0020-7462, DOI: 10.1016/0020-7462(95)00024-0.

— [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.

Nomizu, K. and Sasaki, T. [1994] *Affine Differential Geometry, Geometry of Affine Immersions*, number 111 in Cambridge Tracts in Mathematics, Cambridge University Press: New York/Port Chester/Melbourne/Sydney, ISBN: 978-0-521-44177-3.

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

Vershik, A. M. [1984] *Classical and nonclassical dynamics with constraints*, in *Global Analysis—Studies and Applications. I*, edited by Y. G. Borisovich and Y. E. Gliklikh, 1108 Lecture Notes in Mathematics, pages 278–301, Springer-Verlag: New York/Heidelberg/-Berlin, ISBN: 978-3-540-13910-2.