Affine Grassmannians

Jean Gallier
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
e-mail: jean@cis.upenn.edu

April 28, 2023

Contents

Chapter 1

Affine Grassmannians

1.1 Actions of $\mathbf{SE}(n)$ on Affine Grassmannians

In this section, we show that the Grassmannian $AG(k, n)$ of k-dimensional affine subspaces of $\mathbb{R}^n$ arises as the homogeneous space $\mathbf{SE}(n)/S(\mathbf{E}(k) \times \mathbf{O}(n-k))$, in terms of a transitive action of $\mathbf{SE}(n)$ on $AG(n, k)$.

Recall that a nonempty k-dimensional affine subspace $\mathcal{A}$ of $\mathbb{R}^n$ is determined by a pair (a_0, U), where $a_0 \in \mathbb{R}^n$ is any point in $\mathcal{A}$ and U is a k-dimensional subspace of $\mathbb{R}^n$ called the *direction* of $\mathcal{A}$, with

$$\mathcal{A} = a_0 + U = \{a_0 + u \mid u \in U\}.$$

Two pairs (a_0, U) and (b_0, U) define the same affine subspace $\mathcal{A}$ iff $b_0 - a_0 \in U$ (in fact, U consists of all vectors of the form $b - a$, with $a, b \in \mathcal{A}$).

The subspace U can be represented by any basis $(u_1, \ldots, u_k)$ of vectors $u_i \in U$, and so $\mathcal{A}$ is represented by the *affine frame* $(a_0, (u_1, \ldots, u_k))$.

Two affine frames $(a_0, (u_1, \ldots, u_k))$ and $(b_0, (v_1, \ldots, v_k))$ represent the same affine subspace $\mathcal{A}$ iff there is an invertible $k \times k$ matrix $\Lambda = (\lambda_{ij})$ such that

$$v_j = \sum_{i=1}^{k} \lambda_{ij} u_i, \quad 1 \leq j \leq k,$$

and if there is some vector $c \in \mathbb{R}^k$ such that

$$b_0 = a_0 + \sum_{i=1}^{k} c_i u_i.$$

Note that (Λ, c) defines an invertible affine map of $\mathbb{R}^k$.

A basis $(u_1, \ldots, u_k)$ of U is represented by a $n \times k$ matrix of rank k, say A, so the affine subspace $\mathcal{A}$ is represented by the pair (a_0, A), where $a_0 \in \mathbb{R}^n$ and A is a $n \times k$ matrix of rank

k. The equivalence relation on pairs (a_0, A) is given by

$$(a_0, A) \equiv (b_0, B)$$

iff there exists a pair (Λ, c), where Λ is an invertible $k \times k$ matrix ($\Lambda \in \mathbf{GL}(k, \mathbb{R})$) and c is some vector in $\mathbb{R}^k$, such that

$$B = A\Lambda \quad \text{and} \quad b_0 = a_0 + Ac.$$

Using Gram-Schmidt, we may assume that $(u_1, \ldots, u_k)$ is an orthonormal basis, which means that the columns of the matrix A are orthonormal; that is,

$$A^\top A = I_k.$$

Then, in the equivalence relation defined above, the matrix Λ is an orthogonal $k \times k$ matrix ($\Lambda \in \mathbf{O}(k)$).

Definition 1.1. The (real) *affine Grassmannian* $AG(k, n)$ consists of all k-dimensional affine subspaces of $\mathbb{R}^n$ ($1 \leq k \leq n$).

In the special case $k = 1$, the affine Grassmannian $AG(1, n)$ consists of all affine lines in $\mathbb{R}^n$. This is already a topologically complicated space (more complicated than projective space $\mathbb{RP}^n$).

The (linear) Grassmannian $G(k, n)$ consists of all k-dimensional (linear) subspaces of $\mathbb{R}^n$ ($1 \leq k \leq n$). By linear duality between a finite-dimensional vector space and its dual, $G(k, n)$ is isomorphic to $G(n - k, n)$.

There is a relationship between the affine Grassmannians and the linear Grassmannians. Indeed, we have

$$AG(k, n) = G(k + 1, n + 1) - G(k + 1, n).$$

This is because $G(k + 1, n + 1)$ corresponds to the projective subspaces of dimension k in $\mathbb{RP}^n$. In $\mathbb{R}^{n+1}$, there is a bijection between the set $G(k + 1, n + 1) - G(k + 1, n)$ of linear subspaces V of dimension $k + 1$ that are not contained in the hyperplane of equation $x_{n+1} = 0$, and the set $AG(k, n)$ of k-dimensional affine subspaces of $\mathbb{R}^n$, given by

$$V \mapsto V \cap H_1,$$

where H_1 is the affine hyperplane in $\mathbb{R}^{n+1}$ of equation $x_{n+1} = 1$. The $(k+1)$-dimensional linear subspaces contained in the hyperplane $x_{n+1} = 0$ correspond to the k-dimensional projective subspaces of $\mathbb{RP}^n$ "at infinity" (if we choose the hyperplane $x_{n+1} = 0$ as the hyperplane at infinity in $\mathbb{RP}^n$). As a consequence of the equation $AG(k, n) = G(k+1, n+1) - G(k+1, n)$, the space $AG(k, n)$ is an open subspace of the set of k-dimensional projective subspaces of $\mathbb{RP}^n$, and thus is not compact. Observe that if $0 \leq k \leq n - 1$, then

$$\begin{aligned} A(n - k - 1, n) &= G(n - k, n + 1) - G(n - k, n) \\ &\cong G(k + 1, n + 1) - G(k, n), \end{aligned}$$

so $A(k,n)$ is not isomorphic to $A(n-k-1,n)$, except in the trivial case where $n=2k+1$.

When $n=2$ and $k=1$, we have

$$AG(1,2)=G(2,3)-G(2,2)\cong G(1,3)-G(0,2)=\mathbb{RP}^2-\{\text{one point}\},$$

so $AG(1,2)$ is homeomorphic to the result of deleting one point from the projective plane $\mathbb{RP}^2$, a space homeomorphic to an open Möbius strip (a Möbius strip with its boundary removed). No wonder $AG(1,2)$ is hard to deal with!

Recall that the *Euclidean group* $\mathbf{E}(n)$ consists of all invertible affine maps (Q,u), with $Q\in\mathbf{O}(n)$ and $u\in\mathbb{R}^n$, and that the *special Euclidean group* $\mathbf{SE}(n)$ consists of all invertible affine maps (Q,u), with $Q\in\mathbf{SO}(n)$ and $u\in\mathbb{R}^n$. As usual, we represent an element (Q,u) of $\mathbf{E}(n)$ (or $\mathbf{SE}(n)$) by the $(n+1)\times(n+1)$ matrix

$$\begin{pmatrix} Q & u\\ 0 & 1\end{pmatrix},$$

with $\mathbb{R}^n$ embedded in $\mathbb{R}^{n+1}$ by adding 1 as $(n+1)$th coordinate.

Definition 1.2. Define an action of the group $\mathbf{SE}(n)$ on $AG(k,n)$ as follows: if $\mathcal{A}\in AG(k,n)$, for any affine frame (a_0,A) representing $\mathcal{A}$ (where $A^\top A=I_k$), for any $(Q,u)\in\mathbf{SE}(n)$, then

$$(Q,u)\cdot\mathcal{A}=(Qa_0+u,QA).$$

We need to check that the above action does not depend on the affine frame (a_0,A) chosen for $\mathcal{A}$. If (b_0,B) is another affine frame of $\mathcal{A}$, then there is some orthogonal matrix $\Lambda\in\mathbf{O}(k)$ and some vector $c\in\mathbb{R}^k$ such that

$$B=A\Lambda\quad\text{and}\quad b_0=a_0+Ac,$$

so we have

$$Qb_0=Qa_0+QAc,$$

and

$$QB=QA\Lambda,$$

which shows that (Qa_0+u,QA) and (Qb_0+u,QB) are equivalent *via* (Λ,c), since $QB=(QA)\Lambda$ and $Qb_0+u=Qa_0+u+(QA)c$. Therefore, the action of $\mathbf{SE}(n)$ on $AG(k,n)$ defined above does not depend on the affine frame chosen in $\mathcal{A}$.

The above action is transitive.

Indeed, if (a_0,A) and (b_0,B) represent two affine subspaces, where $A^\top A=I_k$ and $B^\top B=I_k$, then by Gram-Schmidt, we can extend the columns of A into an orthonormal basis A' of $\mathbb{R}^n$, and similarly we can can extend the columns of B into an orthonormal basis B' of $\mathbb{R}^n$. But then, the matrices A' and B' are $n\times n$ orthogonal matrices, and by changing the sign of the their first column if necessary, we may assume that $\det(A')=\det(B')=1$, where the first k columns of A' still define the same subspace as the k columns of A, since they are obtained by multiplying on the right by the $k \times k$ orthogonal matrix $\mathrm{diag}(-1, 1, \ldots, 1)$ (and similarly for B and B'). Write $A'[1..k]$ for the first k columns of A' (and similarly for B'). If we let $Q = B'(A')^\top$ and $u = b_0 - Qa_0$, we have $(Q, u) \in \mathbf{SE}(n)$, and

$$(Q, u) \cdot (a_0, A'[1..k]) = (Qa_0 + b_0 - Qa_0, QA'[1..k]) = (b_0, B'[1..k]);$$

this is because

$$A' = \begin{pmatrix} A_1 & A_2 \end{pmatrix}, \quad B' = \begin{pmatrix} B_1 & B_2 \end{pmatrix},$$

and since A' is orthogonal (so is B'), $A_2^\top A_1 = 0$ and $A_1^\top A_1 = I_k$, so we have

$$\begin{aligned} QA'[1..k] &= B'(A')^\top A'[1..k] \\ &= \begin{pmatrix} B_1 & B_2 \end{pmatrix} \begin{pmatrix} A_1^\top \\ A_2^\top \end{pmatrix} A_1 \\ &= \begin{pmatrix} B_1 A_1^\top + B_2 A_2^\top \end{pmatrix} A_1 \\ &= B_1 A_1^\top A_1 + B_2 A_2^\top A_1 \\ &= B_1 = B'[1..k]. \end{aligned}$$

Therefore, our action is transitive.

Next, we determine the stabilizer of the affine subspace defined by the affine frame $(0, (e_1, \ldots, e_k))$, where $e_1, \ldots, e_k$ are the first k canonical basis vectors of $\mathbb{R}^n$. This affine subspace is also represented by $(0, P_{n,k})$, where $P_{n,k}$ is the $n \times k$ matrix consisting of the first k columns of the identity matrix I_n; namely

$$P_{n,k} = \begin{pmatrix} I_k \\ 0_{n-k,k} \end{pmatrix}.$$

Proposition 1.1. *The stabilizer of the affine subspace defined by* $(0, P_{n,k})$ *is the group* $H = S(\mathbf{E}(k) \times \mathbf{O}(n-k))$ *given by the set of matrices*

$$H = \left\{ \begin{pmatrix} Q & 0 & u \\ 0 & R & 0 \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; Q \in \mathbf{O}(k),\ R \in \mathbf{O}(n-k),\ \det(Q)\det(R) = 1,\ u \in \mathbb{R}^k \right\}.$$

Proof. For any $(P, z) \in \mathbf{SE}(n)$, we have

$$(P, z) \cdot (0, P_{n,k}) = (P0 + z, PP_{n,k}) = (z, P[1..k]).$$

In order for $(z, P[1..k])$ to represent the same affine subspace as $(0, P_{n,k})$, there must be some pair (Λ, c) where $\Lambda \in \mathbf{O}(k)$ and $c \in \mathbb{R}^k$, so that

$$P[1..k] = P_{n,k}\Lambda \quad \text{and} \quad z = P_{n,k}c.$$

The vector $P_{n,k}c$ is obtained from c by adding 0 as the last $n-k$ coordinates, and the matrix $P_{n,k}\Lambda$ is obtained from Λ by adding $n-k$ rows consisting of the vector $\underbrace{(0,\ldots,0)}_{k}$. Therefore, the last $n-k$ coordinates of z must be zero, and the last $n-k$ rows of $P[1..k]$ must be zero rows. Since P is an orthogonal matrix, it must be of the form

$$P = \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix},$$

with $Q \in \mathbf{O}(k)$ and $R \in \mathbf{O}(n-k)$. Since $\det(P) = 1$, we must have $\det(P) = \det(Q)\det(R) = 1$, and the proposition follows. □

1.2 The Grassmannian $AG(k, n)$ as a Reductive Homogeneous Space

In this section, we show that the affine Grassmannian $AG(k, n)$ is a reductive homogeneous space with a simple reductive decomposition $\mathfrak{se}(n) = \mathfrak{h} \oplus \mathfrak{m}$. In fact, there is an involutive automorphism σ of $\mathbf{SE}(n)$ whose fixed subgroup is exactly the group $H = S(\mathbf{E}(k) \times \mathbf{O}(n-k))$ introduced in the previous section. It follows that, except for the fact that there is no $\mathrm{Ad}(H)$-invariant metric on $\mathfrak{m}$ (because H is not compact), all the other properties of a symmetric space are satisfied.

Let $I_{k,n-k}$ be the matrix

$$I_{k,n-k} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix}.$$

Note that $I_{k,n-k}^2 = I_n$. We define an automorphism σ of $\mathbf{SE}(n)$ as follows:

$$\sigma \begin{pmatrix} Q & z \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Q & z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix}.$$

Because $I_{k,n-k}^2 = I_n$, we have $\sigma^2 = \mathrm{id}$. Let us find the subgroup $\mathbf{SE}(n)^\sigma$ of $\mathbf{SE}(n)$ fixed by σ. Every matrix P in $\mathbf{SE}(n)$ can be written as

$$P = \begin{pmatrix} Q & R & u \\ S & T & v \\ 0 & 0 & 1 \end{pmatrix},$$

with $u \in \mathbb{R}^k$ and $v \in \mathbb{R}^{n-k}$, and we have

$$\begin{aligned} \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_{n-k} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} Q & R & u \\ S & T & v \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_{n-k} & 0 \\ 0 & 0 & 1 \end{pmatrix} &= \begin{pmatrix} Q & R & u \\ -S & -T & -v \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_{n-k} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} Q & -R & u \\ -S & T & -v \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Then, $\sigma(P) = P$ iff

$$R = -R, \quad S = -S, \quad v = -v,$$

which means that $R = 0$, $S = 0$, and $v = 0$. Therefore $\mathbf{SE}(n)^\sigma = S(\mathbf{E}(k) \times \mathbf{O}(n-k)) = H$.

The Lie algebras of $\mathbf{SE}(n)$ and $H = \mathbf{SE}(n)^\sigma$ are

$$\mathfrak{se}(n) = \left\{ \begin{pmatrix} S & -A^\top & u \\ A & T & v \\ 0 & 0 & 0 \end{pmatrix} \,\middle|\, S \in \mathfrak{so}(k),\ T \in \mathfrak{so}(n-k),\ A \in \mathrm{M}_{n-k,k},\ u \in \mathbb{R}^k,\ v \in \mathbb{R}^{n-k} \right\}$$

and

$$\mathfrak{h} = \left\{ \begin{pmatrix} S & 0 & u \\ 0 & T & 0 \\ 0 & 0 & 0 \end{pmatrix} \,\middle|\, S \in \mathfrak{so}(k),\ T \in \mathfrak{so}(n-k),\ u \in \mathbb{R}^k \right\}.$$

The derivative $\theta = d\sigma_I$ is an involutive automorphism of $\mathfrak{se}(n)$ which is easily found using curves through I. For any $X \in \mathfrak{se}(n)$, if γ is the curve in $\mathbf{SE}(n)$ given by $\gamma(t) = e^{tX}$, then $\gamma(0) = I$, $\gamma'(0) = X$, and by the chain rule

$$\left.\frac{d(\sigma(\gamma(t))}{dt}\right|_{t=0} = d\sigma_{\gamma(0)}(\gamma'(0)) = d\sigma_I(X),$$

so we have

$$\begin{aligned} d\sigma_I(X) &= \frac{d}{dt}\left(\begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix} e^{tX} \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix}\right)_{t=0} \\ &= \left(\begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix} X e^{tX} \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix}\right)_{t=0} \\ &= \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix} X \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Therefore

$$\theta\begin{pmatrix} S & z \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S & z \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_{k,n-k} & 0 \\ 0 & 1 \end{pmatrix}.$$

Consequently, the Lie algebra $\mathfrak{h}$ is the eigenspace of θ associated with the eigenvalue $+1$, whereas the eigenspace of θ associated with the eigenvalue -1 is given by

$$\mathfrak{m} = \left\{ \begin{pmatrix} 0 & -A^\top & 0 \\ A & 0 & v \\ 0 & 0 & 0 \end{pmatrix} \,\middle|\, A \in \mathrm{M}_{n-k,k},\ v \in \mathbb{R}^{n-k} \right\}.$$

By Lemma 30 in O'Neill [4] (Chapter 11), the fact that σ is an involutive automorphism of $\mathbf{SE}(n)$ whose fixed subgroup is H has the following interesting implications.

Proposition 1.2. *The following properties hold:*

(1) We have a direct sum

$$\mathfrak{se}(n) = \mathfrak{h} \oplus \mathfrak{m}.$$

(2) $\mathrm{Ad}(h)(\mathfrak{m}) \subseteq \mathfrak{m}$ *for all* $h \in H$.

(3) We have

$$[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}, \quad [\mathfrak{h}, \mathfrak{m}] \subseteq \mathfrak{m}, \quad [\mathfrak{m}, \mathfrak{m}] \subseteq \mathfrak{h}.$$

Consequently, $AG(k, n)$ *is a reductive homogeneous space with the reductive decomposition* $\mathfrak{se}(n) = \mathfrak{h} \oplus \mathfrak{m}$.

The next step is to check whether it is possible to define a G-invariant metric on $AG(k, n)$. For this, let us figure out what the adjoint action of H on $\mathfrak{m}$ is. For any

$$h = \begin{pmatrix} R & 0 & u \\ 0 & S & 0 \\ 0 & 0 & 1 \end{pmatrix} \in H,$$

and any

$$X = \begin{pmatrix} 0 & -A^\top & 0 \\ A & 0 & v \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{m},$$

we have

$$\begin{aligned} \mathrm{Ad}_h(X) &= \begin{pmatrix} R & 0 & u \\ 0 & S & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -A^\top & 0 \\ A & 0 & v \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} R^\top & 0 & -R^\top u \\ 0 & S^\top & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -RA^\top & 0 \\ SA & 0 & Sv \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} R^\top & 0 & -R^\top u \\ 0 & S^\top & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -RA^\top S^\top & 0 \\ SAR^\top & 0 & -SAR^\top u + Sv \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Consider the matrices $h \in H$ such that $R = I, S = I$ and the first coordinate u_1 in u is nonzero. The matrices $X \in \mathfrak{m}$ that either have a single nonzero entry equal to 1 in A (and $A^\top$) or a single nonzero entry in v form a basis of $\mathfrak{m}$. Let $E_{k+11} \in \mathfrak{m}$ be the matrix whose only nonzero entries are $e_{k+11} = 1$ and $e_{1k+1} = -1$, and let $E_{k+1n} \in \mathfrak{m}$ be the matrix whose only nonzero entry is $e_{k+1n} = 1$. Then, we have

$$\mathrm{Ad}_h(E_{k+11}) = \begin{pmatrix} 0 & A^\top & 0 \\ A & 0 & -Au \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & \cdots & 0 & 0 & 0 & \cdots & -u_1 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} = E_{k+11} - u_1 E_{k+1n}.$$

Therefore, the matrix of Ad_h over the basis (E_{ij}) has the entry $-u_1$ in the row corresponding to E_{k+1n} and the column corresponding to E_{k+11}, and since $u_1 \in \mathbb{R}$ is arbitrary, we see that the matrices representing the linear maps Ad_h have unbounded entries (even for the special kinds of matrices in h that we are considering). Therefore, $\mathrm{Ad}(H)$ is not bounded, and thus it closure is not compact, which implies that there is no $\mathrm{Ad}(H)$-invariant inner product on $\mathfrak{m}$ (by Theorem 2.42 of Gallot, Hullin, Lafontaine [1]). Therefore, there is no hope for a G-invariant metric on $AG(k, n)$. Except for that, $AG(k, n)$ has all the other properties of a symmetric space.

1.3 Metrics on G/H Induced by Right-Invariant Metrics on G

Given a reductive homogeneous manifold G/H with reductive decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, if H is not compact, $\mathrm{Ad}(H)$-invariant metrics on $\mathfrak{m}$ do not necessarily exist. It is still desirable to obtain metrics on G/H such that the projection $\pi\colon G \to G/H$ is a Riemannian submersion. Since H acts freely and properly on G *on the right*, for every right-invariant metric on G induced by an inner product $\langle -, - \rangle$ on $\mathfrak{g}$, the maps R_h are isometries for all $h \in H$, so by Proposition 2.28 in Gallot, Hullin, Lafontaine [1], (Chapter 2), there is a unique Riemannian metric $\langle -, - \rangle_{G/H}$ on G/H such that $\pi\colon G \to G/H$ is a Riemannian submersion.

Since G/H is reductive, we have

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m},$$

and this makes it possible to pick the horizontal subspaces in the tangent spaces T_aG (with $a \in G$) in terms of $\mathfrak{m}$ and to give a more direct proof of Proposition 2.28 from Gallot, Hullin, Lafontaine [1].

Given an inner product $\langle -, - \rangle$ on $\mathfrak{g}$, recall that the induced right-invariant metric on G is given by

$$\langle u, v \rangle_a = \langle (dR_{a^{-1}})_a(u), dR_{a^{-1}})_a(v) \rangle, \quad \text{for all } u, v \in T_aG \text{ and all } a \in G.$$

We will show that a metric on G/H can be obtained by propagating by right-invariance a metric on $\mathfrak{g}$ to all of the "horizontal subspaces" $(dL_a)_1(\mathfrak{m})$ of $T_a(G) = (dL_a)_1(\mathfrak{g})$.

Because of the invariance condition $\mathrm{Ad}_h(\mathfrak{m}) \subseteq \mathfrak{m}$ for all $h \in H$ (since G/H is a reductive homogeneous space), since $\mathfrak{m}$ is finite-dimensional and Ad_h is injective, we have $\mathrm{Ad}_h(\mathfrak{m}) = \mathfrak{m}$, and if $b = ah$ then

$$\mathrm{Ad}_b = \mathrm{Ad}_a \circ \mathrm{Ad}_h,$$

which implies that

$$\mathrm{Ad}_b(\mathfrak{m}) = \mathrm{Ad}_a(\mathfrak{m}), \quad \text{for all } a, b \in G \text{ such that } a^{-1}b \in H.$$

This means that $\mathrm{Ad}_a(\mathfrak{m})$ depends only on the point $p \in G/H$ for which $p = aH$.

Recall that for every $a \in G$, the map $\tau_a \colon G/H \to G/H$ is defined by

$$\tau_a(bH) = abH, \quad \text{for all } a, b \in G,$$

and $\pi \colon G \to G/H$ is the projection given by $\pi(a) = aH$. For all $a, b \in G$, we have

$$\tau_a(\pi(b)) = abH = \pi(L_a(b)),$$

namely

$$\tau_a \circ \pi = \pi \circ L_a.$$

By taking the derivative at 1, we get

$$(d\tau_a)_o \circ d\pi_1 = d\pi_a \circ (dL_a)_1;$$

equivalently, the following diagram commutes (where $p = aH$):

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{(dL_a)_1} & (dL_a)_1(\mathfrak{g}) = T_aG \\ \Big\downarrow{\scriptstyle d\pi_1} & & \Big\downarrow{\scriptstyle d\pi_a} \\ T_o(G/H) & \xrightarrow[(d\tau_a)_o]{} & T_p(G/H). \end{array}$$

Since $\mathrm{Ker}\, d\pi_1 = \mathfrak{h}$ and since $(dL_a)_1$ is an isomorphism, we see that

$$\mathrm{Ker}\, d\pi_a = (dL_a)_1(\mathfrak{h}).$$

Also, since the restriction of $d\pi_1$ to $\mathfrak{m}$ is an isomorphism and $(dL_a)_1$ and $(d\tau_a)_o$ are isomorphims, so is the restriction of $d\pi_a$ to $(dL_a)_1(\mathfrak{m})$. We have the following commutative diagram in which all the maps are isomorphisms (with $p = aH$):

$$\begin{array}{ccc} \mathfrak{m} & \xrightarrow{(dL_a)_1} & (dL_a)_1(\mathfrak{m}) \\ \Big\downarrow{\scriptstyle d\pi_1} & & \Big\downarrow{\scriptstyle d\pi_a} \\ T_o(G/H) & \xrightarrow[(d\tau_a)_o]{} & T_p(G/H). \end{array}$$

For all $a \in G$ and all $h \in H$, we have

$$\pi(a) = aH = ahH = \pi(ah) = \pi(R_h(a));$$

that is, $\pi = \pi \circ R_h$, and by taking derivatives at a, we get

$$d\pi_a = d\pi_{ah} \circ d(R_h)_a. \tag{$*$}$$

Equivalently, if we write $b = ah$ and $p = aH$ for $a \in G$ and $h \in H$, we have the following commutative diagram:

$$\begin{array}{ccc} T_aG = (dL_a)_1(\mathfrak{g}) & \xrightarrow{(dR_h)_a} & (dL_b)_1(\mathfrak{g}) = T_bG \\ & {\scriptstyle d\pi_a} \searrow \quad \swarrow {\scriptstyle d\pi_b} & \\ & T_p(G/H) & \end{array}$$

Since

$$T_aG = (dL_a)_1(\mathfrak{g}),$$

we have

$$T_aG = (dL_a)_1(\mathfrak{h}) \oplus (dL_a)_1(\mathfrak{m}),$$

and since $\mathrm{Ker}\, d\pi_a = (dL_a)_1(\mathfrak{h})$ and the restriction of $d\pi_a$ to $(dL_a)_1(\mathfrak{m})$ is an isomorphism onto $T_p(G/H)$, we can take $(dL_a)_1(\mathfrak{m})$ as the horizontal subspace $\mathcal{H}_a$ of T_aG.

As a consequence, for any $p \in G/H$ and any $a \in G$ such that $p = aH$, since the map $d\pi_a\colon (dL_a)_1(\mathfrak{m}) \to T_p(G/H)$ is an isomorphism, for any $u \in T_p(G/H)$, there is a unique $X \in \mathfrak{m}$ such that

$$u = (d\pi_a \circ (dL_a)_1)(X);$$

namely, $X = ((dL_{a^{-1}})_a \circ (d\pi_a)^{-1})(u)$.

Let us find out how X changes when we express u in terms of b, with $b = ah$ for some $h \in H$.

Proposition 1.3. *For any $p = aH = bH$ in G/H, if $b = ah$ for some $h \in H$, for any $u \in T_p(G/H)$ and any $X \in \mathfrak{m}$ such that $u = (d\pi_a \circ (dL_a)_1)(X)$, we have*

$$u = (d\pi_b \circ (dL_b)_1)(X'), \quad \textit{with} \quad X' = \mathrm{Ad}_{h^{-1}}(X).$$

Proof. Since $a = bh^{-1}$, by $(*)$ $d\pi_a = d\pi_b \circ d(R_h)_a$, L_a and R_h commute, and since $\mathrm{Ad}_a = d(L_a \circ R_{a^{-1}})_1 = (dL_a)_{a^{-1}} \circ (dR_{a^{-1}})_1$, we have

$$\begin{aligned} u &= (d\pi_a \circ (dL_a)_1)(X) \\ &= (d\pi_b \circ (dR_h)_a \circ (dL_a)_1)(X) \\ &= (d\pi_b \circ (dL_a)_h \circ (dR_h)_1)(X) \\ &= (d\pi_b \circ (dL_b)_1 \circ (dL_{h^{-1}})_h \circ (dR_h)_1)(X) \\ &= (d\pi_b \circ (dL_b)_1 \circ \mathrm{Ad}_{h^{-1}})(X), \end{aligned}$$

which shows that

$$u = (d\pi_b \circ (dL_b)_1)(X'), \quad \text{with} \quad X' = \mathrm{Ad}_{h^{-1}}(X),$$

as claimed. □

For any $a \in G$, the map $\mathrm{Ad}_a : \mathfrak{g} \to \mathfrak{g}$ is a linear isomorphism of $\mathfrak{g}$, so $\mathrm{Ad}_a(\mathfrak{m})$ is always a subspace of $\mathfrak{g}$. In the special case where $a \in H$, we have $\mathrm{Ad}_a(\mathfrak{m}) = \mathfrak{m}$, but for $a \in G - H$, this is generally false and we can only claim that $\mathrm{Ad}_a(\mathfrak{m}) \subseteq \mathfrak{g}$. Here is the main theorem of this section.

Theorem 1.4. *Given any homogeneous reductive manifold G/H with reductive decomposition*

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m},$$

every inner product $\langle -, - \rangle$ on $\mathfrak{g}$ yields a Riemannian metric on G/H such that if G is endowed with the right-invariant Riemannian metric induced by $\langle -, - \rangle$, then $\pi : G \to G/H$ is a Riemannian submersion. For every $a \in G$, the horizontal subspace $\mathcal{H}_a$ at a is given by

$$\mathcal{H}_a = (dL_a)_1(\mathfrak{m}),$$

and the restriction of $d\pi_a$ to $\mathcal{H}_a = (dL_a)_1(\mathfrak{m})$ is an isomorphism onto $T_p(G/H)$, with $p = aH$. The metric on $T_p(G/H)$ is defined as follows: For every $p = aH \in G/H$, for any two vectors $u, v \in T_p(G/H)$,

$$\langle u, v \rangle_{G/H,p} = \langle (d\pi_a)^{-1}(u), (d\pi_a)^{-1}(v) \rangle_a,$$

where $\langle -, - \rangle_a$ is the right-invariant metric on $\mathcal{H}_a = (dL_a)_1(\mathfrak{m})$ induced by the inner product on $\mathfrak{g}$, which means that

$$\langle u, v \rangle_{G/H,p} = \langle (dR_{a^{-1}})_a((d\pi_a)^{-1}(u)), (dR_{a^{-1}})_a((d\pi_a)^{-1}(v)) \rangle.$$

Equivalently, if X and Y are the unique vectors in $\mathfrak{m}$ such that $X = ((dL_{a^{-1}})_a \circ (d\pi_a)^{-1})(u)$ and $Y = ((dL_{a^{-1}})_a \circ (d\pi_a)^{-1})(v)$, then

$$\langle u, v \rangle_{G/H,p} = \langle \mathrm{Ad}_a(X), \mathrm{Ad}_a(Y) \rangle.$$

Furthermore, $\mathrm{Ad}_a(\mathfrak{m})$ depends only on the point $p \in G/H$ for which $p = aH$. We can choose an inner product on $\mathfrak{g}$ by picking any inner product on $\mathfrak{m}$ and any inner product on $\mathfrak{h}$ and asserting that $\mathfrak{h}$ and $\mathfrak{m}$ are orthogonal. This, way

$$T_aG = (dL_a)_1(\mathfrak{h}) \oplus (dL_a)_1(\mathfrak{m}),$$

where the vertical subspace $\mathcal{V}_a = (dL_a)_1(\mathfrak{h})$ and the horizontal subspace $\mathcal{H}_a = (dL_a)_1(\mathfrak{m})$ are orthogonal for every $a \in G$. Furthermore, for all $a, b \in G$ and all $h \in H$, if $b = ah$, then $(dR_h)_a$ is an isometry between $\mathcal{H}_a$ and $\mathcal{H}_b$.

Proof. We define the metric $\langle -,-\rangle_{G/H}$ using the isomorphisms $d\pi_a\colon \mathcal{H}_a \to T_p(G/H)$, where $\mathcal{H}_a = (dL_a)_1(\mathfrak{m})$, $a \in G$, and $p = aH \in G/H$, as follows. For any two vectors $u, v \in T_p(G/H)$, if X and Y are the unique vectors in $\mathfrak{m}$ such that $X = ((dL_{a^{-1}})_a \circ (d\pi_a)^{-1})(u)$ and $Y = ((dL_{a^{-1}})_a \circ (d\pi_a)^{-1})(v)$, then we have

$$\begin{aligned}\langle u, v\rangle_{G/H,p} &= \langle (dR_{a^{-1}})_a((d\pi_a)^{-1}(u)), (dR_{a^{-1}})_a((d\pi_a)^{-1}(v))\rangle \\ &= \langle (dR_{a^{-1}})_a((dL_a)_1(X)), (dR_{a^{-1}})_a((dL_a)_1(Y))\rangle \\ &= \langle \mathrm{Ad}_a(X), \mathrm{Ad}_a(Y)\rangle.\end{aligned}$$

Thus, the metric on $T_p(G/H)$ is completely determined by the metric on $\mathrm{Ad}_a(\mathfrak{m})$, a subspace which depends only on the point $p \in G/H$ for which $p = aH$.

Let us check that this definition does not depend on the choice of the coset representative $aH = p$. If $bH = aH$, we have $b = ah$ for some $h \in H$, and then by Proposition 1.3 we have

$$X' = ((dL_{b^{-1}})_b \circ (d\pi_b)^{-1})(u) = \mathrm{Ad}_{h^{-1}}(X) \quad \text{and} \quad Y' = ((dL_{b^{-1}})_b \circ (d\pi_b)^{-1})(v) = \mathrm{Ad}_{h^{-1}}(Y),$$

so for $p = bH$, we have

$$\begin{aligned}\langle u, v\rangle_{G/H,p} &= \langle \mathrm{Ad}_b(X'), \mathrm{Ad}_b(Y')\rangle \\ &= \langle \mathrm{Ad}_b(\mathrm{Ad}_{h^{-1}}(X)), \mathrm{Ad}_b(\mathrm{Ad}_{h^{-1}}(Y))\rangle \\ &= \langle \mathrm{Ad}_a(X), \mathrm{Ad}_a(Y)\rangle,\end{aligned}$$

proving that the definition of $\langle u, v\rangle_{G/H,p}$ does not depend on the coset representative of $p = aH$. The smoothness of this metric follows from the standard argument; namely, G is a principal H-bundle over G/H, and so local sections exist.

Observe that the definition

$$\langle u, v\rangle_{G/H,p} = \langle (dR_{a^{-1}})_a((d\pi_a)^{-1}(u)), (dR_{a^{-1}})_a((d\pi_a)^{-1}(v))\rangle$$

means that

$$\langle u, v\rangle_{G/H,p} = \langle (d\pi_a)^{-1}(u), (d\pi_a)^{-1}(v)\rangle_a,$$

where $\langle -,-\rangle_a$ is the right-invariant metric on $\mathcal{H}_a = (dL_a)_1(\mathfrak{m})$ induced by the inner product on $\mathfrak{g}$. Consequently, for all $p \in G/H$ and for all $a \in G$ such that $p = aH$, the isomorphism $d\pi_a\colon \mathcal{H}_A \to T_p(G/H)$ is an isometry, which shows that the submersion π is a Riemannian submersion. Furthermore, for all $a, b \in G$ and all $h \in H$, if $b = ah$, then $(dR_h)_a$ is an isometry between $\mathcal{H}_a = (dL_a)_1(\mathfrak{m})$ and $\mathcal{H}_b = (dL_b)_1(\mathfrak{m})$. ☐

1.4 Connections on Reductive Homogeneous Spaces

Given a reductive homogeneous space G/H with reductive decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, we know that there is a one–to–one correspondence between G-invariant metrics on G/H and inner products $\langle -,\rangle_{\mathfrak{m}}$ on $\mathfrak{m}$ that are $\mathrm{Ad}(H)$-*invariant*, which means that

$$\langle u, v\rangle_{\mathfrak{m}} = \langle \mathrm{Ad}_h(u), \mathrm{Ad}_h(v)\rangle_{\mathfrak{m}}, \quad \text{for all } h \in H \text{ and all } u, v \in \mathfrak{m}.$$

Unfortunately, if H is not compact, such inner products do not exist.

Instead of trying to define a connection on G/H in terms of a metric, we may try to define a connection on G/H in terms of a bilinear map $\alpha\colon \mathfrak{m}\times\mathfrak{m}\to\mathfrak{m}$ on $\mathfrak{m}$. Since the Levi–Civita connection is invariant under diffeomorphisms, the Levi–Civita connection induced by an $\mathrm{Ad}(H)$-invariant inner product on $\mathfrak{m}$ is G-invariant, so it it natural to look for G-invariant connections. Let us review what it means for a connection on G/H to be G-invariant.

Every group element $a\in G$ defines a diffeomorphism $\tau_a\colon G/H\to G/H$ given by

$$\tau_a(gH) = agH, \quad \text{for all } g\in G.$$

Oberve that

$$\tau_{ab} = \tau_a\circ\tau_b,$$

since

$$\tau_{ab}(gH) = abgH = \tau_a(bgH) = \tau_a(\tau_b(gH)),$$

and

$$\tau_h(H) = H, \quad \text{for all } h\in H.$$

Given a diffeomorphism $\varphi\colon M\to N$ between two manifolds M and N, for any vector field V on M, recall that we define the *push-forward* φ_*V of V by

$$(\varphi_*V)_{\varphi(p)} = d\varphi_p V_p, \quad \text{for all } p\in M.$$

If ψ is a diffeomorphism from N to P, then

$$\begin{aligned}((\psi\circ\varphi)_*V)_{\psi(\varphi(p))} &= d(\psi\circ\varphi)_p V_p\\ &= d\psi_{\varphi(p)}(d\varphi_p V_p)\\ &= d\psi_{\varphi(p)}(\varphi_*V)_{\varphi(p)}\\ &= (\psi_*(\varphi_*V))_{\psi(\varphi(p))},\end{aligned}$$

which shows that

$$(\psi\circ\varphi)_* = \psi_*\circ\varphi_*.$$

Definition 1.3. A connection ∇ on a homogeneous space G/H is *G-invariant* if

$$(\tau_a)_*(\nabla_V W) = \nabla_{(\tau_a)_*V}((\tau_a)_*W), \quad \text{for all } V,W\in\mathcal{X}(G/H) \text{ and all } a\in G. \qquad (*)$$

Recall that $(\nabla_V W)_p$ depend only of V_p, so

$$(\nabla_V W)_p = (\nabla_{V_p} W)_p.$$

We make constant use of the above fact.

The natural projection from G onto G/H is denoted by $\pi\colon G\to G/H$. Recall that the restriction of the map $d\pi_1\colon \mathfrak{g}\to T_0(G/H)$ to $\mathfrak{m}$ is a linear isomorphism (where o denotes the point in G/H corresponding to the coset $1H = H$). Since $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, every vector $X \in \mathfrak{g}$ has a unique decomposition as

$$X = X_{\mathfrak{h}} + X_{\mathfrak{m}}, \quad X_{\mathfrak{h}} \in \mathfrak{h}, X_{\mathfrak{m}} \in \mathfrak{m}.$$

The fact that every $X \in \mathfrak{g}$ induces a vector field X^* on G/H (an *action field* or *infinitesimal generator*) through the left action of G on G/H plays a crucial role. For any $X \in \mathfrak{g}$, the vector field X^* is given by

$$X^*_p = \frac{d}{dt}(\exp(tX)aH)\bigg|_{t=0},$$

for any $a \in G$ such that $p = aH$. Recall that the linear map $d\pi_1 \colon \mathfrak{g} \to T_o(G/H)$ can be expressed as

$$d\pi_1(X) = X^*_0 = \frac{d}{dt}(\exp(tX)H)\bigg|_{t=0},$$

and $\mathrm{Ker}\,(d\pi_1) = \mathfrak{h}$.

It turns out that any G-invariant connection on G/H is uniquely determined by the bilinear map $\alpha \colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ given by

$$\alpha(X, Y) = (d\pi_1)^{-1}(\nabla_{X^*_o} Y^*)_o, \quad \text{for all } X, Y \in \mathfrak{m}.$$

Furhermore, there is a one-to-one correspondence between G-invariant connection on G/H and bilinear maps $\alpha \colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ satifying the condition

$$\mathrm{Ad}_h(\alpha(X, Y)) = \alpha(\mathrm{Ad}_h(X), \mathrm{Ad}_h(Y)), \quad \text{for all } X, Y \in \mathfrak{m} \text{ and all } h \in H. \tag{$\dagger$}$$

It is also possible to characterize torsion-free G-invariant connections and G-invariant connections for which the geodesics through o are of the form $\gamma(t) = e^{tX} \cdot o$. In this case, the bilinear map α is given by

$$\alpha(X, Y) = -\frac{1}{2}[X, Y]_{\mathfrak{m}}, \quad X, Y \in \mathfrak{m}.$$

This connection is known as the *Cartan connection* on G/H. The Levi–Civita connection associated with a G-invariant metric on G/H coincides with the Cartan connection on G/H iff G/H is naturally reductive.

The following technical results will be needed.

Proposition 1.5. *For any $X \in \mathfrak{g}$ and any $a \in G$, we have*

$$(\tau_a)_* X^* = (\mathrm{Ad}_a(X))^*.$$

Proof. By definition, for any $p = bH$, we have $\tau_a(bH) = abH$, and

$$\begin{aligned} ((\tau_a)_* X^*)_{\tau_a(p)} &= (d\tau_a)_p(X^*(p)) \\ &= \frac{d}{dt}(a\exp(tX)bH)\bigg|_{t=0} \\ &= \frac{d}{dt}(a\exp(tX)a^{-1}abH)\bigg|_{t=0} \\ &= \frac{d}{dt}(\exp(t\mathrm{Ad}_a(X))abH)\bigg|_{t=0} \\ &= (\mathrm{Ad}_a(X))^*_{\tau_a(p)}, \end{aligned}$$

which shows that $(\tau_a)_* X^* = (\mathrm{Ad}_a(X))^*$. $\square$

In the special case where $p = o$, since $X_o^* = d\pi_1(X)$ for any $X \in \mathfrak{g}$, the above derivation shows that

$$\begin{aligned} ((\tau_a)_* X^*)_{\tau_a(o)} &= d\tau_a(X_0^*) \\ &= d\tau_a(d\pi_o(X)) \\ &= ((d\tau_a)_o \circ d\pi_1)(X) \\ &= (\mathrm{Ad}_a(X))^*_{\tau_a(o)}, \end{aligned}$$

so $((d\tau_a)_o \circ d\pi_1)(X) = (\mathrm{Ad}_a(X))^*_{\tau_a(o)}$. If we restrict X to belong to $\mathfrak{m}$ and if we let $p = aH = \tau_a(o)$ and define $\eta_a\colon \mathrm{Ad}_a(\mathfrak{m}) \to T_p(G/H)$ by

$$\eta_a(Y) = Y_p^*, \quad Y \in \mathrm{Ad}_a(\mathfrak{m}),$$

then we obtain the following commutative diagram:

$$\begin{array}{ccc} \mathfrak{m} & \xrightarrow{\ \mathrm{Ad}_a\ } & \mathrm{Ad}_a(\mathfrak{m}) \\ {\scriptstyle d\pi_1}\downarrow & & \downarrow{\scriptstyle \eta_a} \\ T_o(G/H) & \xrightarrow[(d\tau_a)_o]{} & T_p(G/H). \end{array} \qquad (**)$$

Since the maps $\mathrm{Ad}_a, d\pi_1$ and $(d\tau_a)_o$ are linear isomorphisms, the map η_a is an isomorphism between $\mathrm{Ad}_a(\mathfrak{m})$ and $T_p(G/H)$, with $p = aH$. Observe that if $h \in H$, then $p = o$, $\mathrm{Ad}_h(\mathfrak{m}) = \mathfrak{m}$, and $\eta_h = d\pi_1$.

Proposition 1.6. *For any $p \in G/H$, for any two coset representatives $bH = aH = p$, if $b = ah$ for some $h \in \mathfrak{m}$, then*

$$\eta_b \circ \mathrm{Ad}_b = \eta_a \circ \mathrm{Ad}_a \circ \mathrm{Ad}_h.$$

Proof. Indeed, by $(**)$ we have

$$\begin{aligned}\eta_a \circ \mathrm{Ad}_a &= (d\tau_a)_o \circ d\pi_1\\ \eta_b \circ \mathrm{Ad}_b &= (d\tau_b)_o \circ d\pi_1\\ d\pi_1 \circ \mathrm{Ad}_h &= (d\tau_h)_o \circ d\pi_1,\end{aligned}$$

and we deduce that

$$\begin{aligned}\eta_b \circ \mathrm{Ad}_b &= (d\tau_b)_o \circ d\pi_1\\ &= (d\tau_a)_o \circ (d\tau_h)_o \circ d\pi_1\\ &= (d\tau_a)_o \circ d\pi_1 \circ \mathrm{Ad}_h\\ &= \eta_a \circ \mathrm{Ad}_a \circ \mathrm{Ad}_h,\end{aligned}$$

as claimed. ☐

We begin with a necessary condition for a connection on G/H to be G-invariant. Recall that as a special case of $(**)$, we have

$$d\pi_1 \circ \mathrm{Ad}_h = (d\tau_h)_o \circ d\pi_1 \quad \text{for all } h \in H,$$

which can be expressed as

$$(\mathrm{Ad}_h(X))^*_o = (d\tau_h)_o(X^*_o) \qquad \text{for all } h \in H \text{ and all } X \in \mathfrak{m}.$$

This equation is also shown in O'Neill [4] (Chapter 11, Proposition 22) and Gallier (Proposition 19.16).

If we apply the identity $(*)$ at $\tau_h(o) = o$ to $V = X^*$, $W = Y^*$ with $X, Y \in \mathfrak{m}$, and to $a = h \in H$, we get

$$(\tau_h)_*(\nabla_{X^*}Y^*)_o = (\nabla_{((\tau_h)_*X^*)_o}(\tau_h)_*Y^*)_o,$$

which is equivalent to

$$\begin{aligned}(d\tau_h)_o(\nabla_{X^*_o}Y^*)_o &= (\nabla_{(d\tau_h)_oX^*_o}(\mathrm{Ad}_h(Y))^*)_o\\ &= (\nabla_{(\mathrm{Ad}_h(X))^*_o}(\mathrm{Ad}_h(Y))^*)_o.\end{aligned}$$

Definition 1.4. The bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ is given by

$$\alpha(X, Y) = (d\pi_1)^{-1}(\nabla_{X^*_o}Y^*)_o, \quad \text{for all } X, Y \in \mathfrak{m},$$

where $d\pi_1$ is the isomorphism from $\mathfrak{m}$ onto $T_o(G/H)$. Equivalently, $\alpha(X, Y)$ is determined by

$$\alpha(X, Y)^*_o = (\nabla_{X^*_o}Y^*)_o, \quad \text{for all } X, Y \in \mathfrak{m}.$$

Proposition 1.7. *The bilinear map α associated with a G-invariant connection ∇ on G/H as in Definition 1.4 satisfies the condition*

$$\mathrm{Ad}_h(\alpha(X,Y)) = \alpha(\mathrm{Ad}_h(X), \mathrm{Ad}_h(Y)), \quad \textit{for all } X, Y \in \mathfrak{m} \textit{ and all } h \in H. \qquad (\dagger)$$

Proof. The equation

$$(d\tau_h)_o(\nabla_{X^*_o} Y^*)_o = (\nabla_{(\mathrm{Ad}_h(X))^*_o} (\mathrm{Ad}_h(Y))^*)_o$$

proved earlier shows that

$$(d\tau_h)_o(d\pi_1(\alpha(X,Y))) = d\pi_1(\alpha(\mathrm{Ad}_h(X), \mathrm{Ad}_h(Y))),$$

so

$$d\pi_1(\mathrm{Ad}_h(\alpha(X,Y))) = d\pi_1(\alpha(\mathrm{Ad}_h(X), \mathrm{Ad}_h(Y))),$$

which, since $d\pi_1$ is an isomorphism from $\mathfrak{m}$ onto $T_o(G/H)$, implies the condition

$$\mathrm{Ad}_h(\alpha(X,Y)) = \alpha(\mathrm{Ad}_h(X), \mathrm{Ad}_h(Y)), \quad \text{for all } X, Y \in \mathfrak{m} \text{ and all } h \in H,$$

as claimed. ☐

Here is the main theorem of this section.

Theorem 1.8. *Given any homogeneous reductive space G/H with reductive decomposition*

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m},$$

there is a one-to-one correspondence between G-invariant connection on G/H and bilinear maps $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ satifying the condition

$$\mathrm{Ad}_h(\alpha(X,Y)) = \alpha(\mathrm{Ad}_h(X), \mathrm{Ad}_h(Y)), \quad \textit{for all } X, Y \in \mathfrak{m} \textit{ and all } h \in H. \qquad (\dagger)$$

Given any G-invariant connection ∇ on G/H, the bilinear map α is given by

$$\alpha(X,Y) = (d\pi_1)^{-1}(\nabla_{X^*_o} Y^*)_o, \quad \textit{for all } X, Y \in \mathfrak{m},$$

where $d\pi_1$ is the isomorphism from $\mathfrak{m}$ onto $T_o(G/H)$. Conversely, given a bilinear map α satisfying condition ($\dagger$), the unique G-invariant connection ∇ associated with α is defined as follows. For any $p \in G/H$, for any coset representative $aH = p$ with $a \in G$, the map $\eta_a\colon \mathrm{Ad}_a(\mathfrak{m}) \to T_p(G/H)$ given by

$$\eta_a(Y) = Y^*_p, \quad Y \in \mathrm{Ad}_a(\mathfrak{m}),$$

is a linear isomorphism such that the following diagram commutes:

$$\begin{array}{ccc} \mathfrak{m} & \xrightarrow{\mathrm{Ad}_a} & \mathrm{Ad}_a(\mathfrak{m}) \\ {\scriptstyle d\pi_1}\downarrow & & \downarrow{\scriptstyle \eta_a} \\ T_o(G/H) & \xrightarrow[(d\tau_a)_o]{} & T_p(G/H). \end{array}$$

Then, for any $V \in T_p(G/H)$ and for any vector field W on G/H of the form $W = (\mathrm{Ad}_a(Y))^*$, with $Y \in \mathfrak{m}$, if $X \in \mathfrak{m}$ is the unique vector such that $V = (\eta_a \circ \mathrm{Ad}_a)(X)$, we set

$$(\nabla_V W)_p = (d\tau_a)_o(\nabla_{(d\tau_{a^{-1}})_p(V)}(\tau_{a^{-1}})_* W)_o = (d\tau_a)_o \circ d\pi_1\big(\alpha\big(X, Y\big)\big). \qquad (\dagger\dagger)$$

Furthermore, the G-invariant connection on G/H associated with α is torsion-free iff

$$\alpha(X, Y) - \alpha(Y, X) = -[X, Y]_{\mathfrak{m}}, \quad \textit{for all } X, Y \in \mathfrak{m}.$$

Proof. It was shown in Proposition 1.7 that the bilinear map α associated with a G-invariant connection ∇ on G/H satisfies (†).

Conversely, we show that any bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ satisfying (†) induces a G-invariant connection on G/H.

For any $a \in G$, since τ_a is a diffeomorphism with inverse $\tau_{a^{-1}}$, for any two vector fields V and W over G/H, if the connection $\nabla_V W$ is G-invariant, since

$$(\tau_a)_* \circ (\tau_{a^{-1}})_* = (\tau_a \circ \tau_{a^{-1}})_* = \mathrm{id}_*,$$

we must have

$$\begin{aligned} \nabla_V W &= (\tau_a)_*((\tau_{a^{-1}})_* \nabla_V W) \\ &= (\tau_a)_*(\nabla_{(\tau_{a^{-1}})_* V}(\tau_{a^{-1}})_* W). \end{aligned}$$

At $p = aH$, since $p = \tau_a(o)$, we get

$$\begin{aligned} (\nabla_V W)_p &= (d\tau_a)_o(\nabla_{(\tau_{a^{-1}})_* V}(\tau_{a^{-1}})_* W)_o \\ &= (d\tau_a)_o(\nabla_{(d\tau_{a^{-1}})_p(V_p)}(\tau_{a^{-1}})_* W)_o. \end{aligned}$$

Moreover, $(\nabla_{(d\tau_{a^{-1}})_p(V_p)}(\tau_{a^{-1}})_* W)_o \in T_o(G/H) \cong \mathfrak{m}$, with $(d\tau_{a^{-1}})_p(V_p) \in T_o(G/H)$ and where $(\tau_{a^{-1}})_* W$ is a vector field whose value at o belongs to $T_o(G/H)$. We can pick some chart of G/H at o with domain U, and then we know that over U, the vector field $(\tau_{a^{-1}})_* W$ can be written as

$$\widehat{W} = (\tau_{a^{-1}})_* W = f_1 X_1^* + \cdots + f_n X_n^*,$$

for some basis $(X_1, \ldots, X_n)$ of $\mathfrak{m}$ and for some smooth functions $f_1, \ldots, f_n$ on U. Since $(\tau_{a^{-1}})_* W$ and $\widehat{W}$ agree near o, we have

$$\begin{aligned} \nabla_{(d\tau_{a^{-1}})_p(V_p)}(\tau_{a^{-1}})_* W &= \nabla_{(d\tau_{a^{-1}})_p(V_p)} \widehat{W} \\ &= \sum_{i=1}^n f_i \nabla_{(d\tau_{a^{-1}})_p(V_p)} X_i^* + \sum_{i=1}^n \Big(\big((d\tau_{a^{-1}})_p(V_p)\big) f_i\Big) X_i^*, \end{aligned}$$

(where $\big((d\tau_{a^{-1}})_p(V_p)\big) f_i$ denotes the directional derivative of f_i in the direction $(d\tau_{a^{-1}})_p(V_p)$), which shows that $\nabla_{(d\tau_{a^{-1}})_p(V_p)}(\tau_{a^{-1}})_* W$ is completely determined by the $\nabla_{(d\tau_{a^{-1}})_p(V_p)} X_i^*$, for $i = 1 \ldots, n$.

Given any $p \in G/H$, for any coset representative $aH = p$, recall that we have an isomorphism $\eta_a \colon \mathrm{Ad}_a(\mathfrak{m}) \to T_p(G/H)$, so for any $V \in T_p(G/H)$, there is a unique $X \in \mathfrak{m}$ so that $V = \eta_a(\mathrm{Ad}_a(X))$. Furthermore, we have

$$\begin{aligned}(d\tau_{a^{-1}})_p(V) &= (d\tau_{a^{-1}})_p(\eta_a(\mathrm{Ad}_a(X)))\\ &= (d\tau_{a^{-1}})_p((d\tau_a)_o \circ d\pi_1)(X))\\ &= d\pi_1(X) = X^*_o.\end{aligned}$$

As a consequence, for any $V \in T_p(G/H)$ and for any vector field W on G/H of the form $W = (\mathrm{Ad}_a(Y))^*$ with $Y \in \mathfrak{m}$, since

$$(\tau_{a^{-1}})_*(W) = (\tau_{a^{-1}})_*(\mathrm{Ad}_a(Y))^* = (\mathrm{Ad}_{a^{-1}}(\mathrm{Ad}_a(Y)))^* = Y^*,$$

we have

$$\begin{aligned}(\nabla_{(d\tau_{a^{-1}})_p(V)}(\tau_{a^{-1}})_*W)_o &= (\nabla_{X^*_o}Y^*)_o\\ &= d\pi_1\big(\alpha\big(X,Y\big)\big).\end{aligned}$$

Therefore, for any coset representative $aH = p$ with $a \in G$, for any $V \in T_p(G/H)$ and for any vector field W on G/H of the form $W = (\mathrm{Ad}_a(Y))^*$, with $Y \in \mathfrak{m}$, if $X \in \mathfrak{m}$ is the unique vector such that $V = (\eta_a \circ \mathrm{Ad}_a)(X)$, we set

$$(\nabla_V W)_p = (d\tau_a)_o(\nabla_{(d\tau_{a^{-1}})_p(V)}(\tau_{a^{-1}})_*W)_o = (d\tau_a)_o \circ d\pi_1\big(\alpha\big(X,Y\big)\big). \tag{††}$$

We need to show that the above definition does not depend on the representative of p, so let $b \in G$ such that $aH = bH$. Then, $b = ah$ for some $h \in H$, and we have

$$V = (\eta_a \circ \mathrm{Ad}_a)(X) = (\eta_b \circ \mathrm{Ad}_b)\big(\mathrm{Ad}_{h^{-1}}(X)\big)$$

and

$$W = (\mathrm{Ad}_a(Y))^* = \big(\mathrm{Ad}_b(\mathrm{Ad}_{h^{-1}}(Y)\big)^*.$$

Since $d\pi_1 \circ \mathrm{Ad}_h = (d\tau_h)_o \circ d\pi_1$, we get

$$\begin{aligned}(d\tau_b)_o(\nabla_{(d\tau_{b^{-1}})_p(V)}(\tau_{b^{-1}})_*W)_o &= (d\tau_a)_o \circ (d\tau_h)_o(\nabla_{(d\tau_{b^{-1}})_p(V)}(\tau_{b^{-1}})_*W)_o\\ &= (d\tau_a)_o \circ (d\tau_h)_o \circ d\pi_1\big(\alpha\big(\mathrm{Ad}_{h^{-1}}(X), \mathrm{Ad}_{h^{-1}}(Y)\big)\big)\\ &= (d\tau_a)_o \circ d\pi_1 \circ \mathrm{Ad}_h\big(\alpha\big(\mathrm{Ad}_{h^{-1}}(X), \mathrm{Ad}_{h^{-1}}(Y)\big)\big).\end{aligned}$$

Using (†), this yield

$$\begin{aligned}(d\tau_a)_o \circ d\pi_1 \circ \mathrm{Ad}_h\big(\alpha\big(\mathrm{Ad}_{h^{-1}}(X), \mathrm{Ad}_{h^{-1}}(Y)\big)\big) &= (d\tau_a)_o \circ d\pi_1 \circ \mathrm{Ad}_h \circ \mathrm{Ad}_{h^{-1}}\big(\alpha\big(X,Y\big)\big)\\ &= (d\tau_a)_o \circ d\pi_1\big(\alpha\big(X,Y\big)\big),\end{aligned}$$

which proves that our definition does not depend on the choice of the representative of the coset p. The definition also makes it clear that the resulting connection is G-invariant.

If the connection ∇ is torsion-free, let us find out which condition is imposed on α. Recall that the torsion of a connection ∇ is given by

$$T(V, W) = \nabla_V W - \nabla_W V - [V, W].$$

If the connection ∇ is torsion-free, which means that

$$\nabla_V W - \nabla_W V = [V, W], \quad \text{for all } V, W \in \mathcal{X}(G/H),$$

then we have

$$\nabla_{X^*} Y^* - \nabla_{Y^*} X^* = [X^*, Y^*], \quad \text{for all } X, Y \in \mathfrak{m},$$

which implies that

$$d\pi_1(\alpha(X, Y)) - d\pi_1(\alpha(Y, X)) = -[X, Y]_o^*.$$

However, $[X, Y]_{\mathfrak{m}}$ is the unique vector in $\mathfrak{m}$ such that $d\pi_1([X, Y]_{\mathfrak{m}}) = [X, Y]_o^*$, so we get $d\pi_1(\alpha(X, Y)) - d\pi_1(\alpha(Y, X)) = -d\pi_1([X, Y]_{\mathfrak{m}})$, and since $d\pi_1$ is a bijection from $\mathfrak{m}$ onto $T_o(G/H)$, we obtain

$$\alpha(X, Y) - \alpha(Y, X) = -[X, Y]_{\mathfrak{m}}, \quad \text{for all } X, Y \in \mathfrak{m}.$$

Therefore, if the G-invariant connection ∇ is torsion-free, then $\alpha_S = (\alpha(X, Y) - \alpha(Y, X))/2$, the skew-symmetric part of α, is given by

$$\alpha_S(X, Y) = -\frac{1}{2}[X, Y]_{\mathfrak{m}}.$$

The converse is clear. ☐

Remark: It should be possible to derive Theorem 1.8 from Theorem 2.1 in Kobayashi and Nomizu [3] (Chapter X), a more general result which applies to certain principal subbundles of the bundle of linear frames with structure group some subgroup of $\mathbf{GL}(n, \mathbb{R})$, on a reductive homogeneous space. However, Kobayashi and Nomizu use a different definition of a connection, namely in terms of $\mathfrak{g}$-valued one-forms (so called Ereshmann connections; see Kobayashi and Nomizu [2], Chapters II and III). The translation of their results to connections defined as operators ∇ on vector fields appears to require as much work as proving our theorem directly.

We now find a necessary and sufficient condition on the bilinear map α associated with a G-invariant connection ∇ on G/H so that the curves $\gamma(t) = e^{tX} o = \tau_{e^{tX}}(o)$ through o with $X \in \mathfrak{m}$ are geodesics. Such a condition is given in Kobayashi and Nomizu [3] (Chapter X, Proposition 2.9 and Theorem 2.10). However, as noted earlier, Kobayashi and Nomizu use a different definition of a connection, namely in terms of $\mathfrak{g}$-valued one-forms. The translation of their results to connections defined as operators ∇ on vector fields requires a fair amount of work.

We need a preliminary result. First, observe that for any fixed t, $e^{tX} \in G$ defines the diffeomorphisn $\tau_{e^{tX}}$ of G/H.

Proposition 1.9. *For any reductive homogeneous manifold G/H, for any $X \in \mathfrak{g}$, if γ is the curve in G/H given by $\gamma(t) = e^{tX} \cdot o = \tau_{e^{tX}}(o)$, then for every $t \in \mathbb{R}$, we have*

$$(\tau_{\gamma(t)})_* X^* = X^*.$$

Proof. Since the action vector field X^* is defined such that for any $p \in G/H$,

$$X^*_p = \frac{d}{ds}(e^{sX} aH)\bigg|_{s=0},$$

for any $a \in G$ such that $p = aH$, we have

$$\begin{aligned}(\tau_{\gamma(t)})_* X^*_p &= \frac{d}{ds}(e^{tX} e^{sX} aH)\bigg|_{s=0} \\ &= \frac{d}{ds}(e^{sX} e^{tX} aH)\bigg|_{s=0} \\ &= X^*_{\tau_{\gamma(t)}(p)},\end{aligned}$$

which proves our claim. ∎

Proposition 1.10. *Given any reductive homogeneous manifold G/H and any G-invariant connection ∇ on G/H, for any $X \in \mathfrak{m}$, if γ is the curve in G/H given by $\gamma(t) = e^{tX} \cdot o = \tau_{e^{tX}}(o)$, then γ is a geodesic in G/H iff the bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ associated with ∇ is skew-symmetric (that is, $\alpha(X, X) = 0$ for all $X \in \mathfrak{m}$).*

Proof. (After Kobayashi and Nomizu [3], Proposition 2.9). The curve $\gamma(t) = e^{tX} \cdot o$ is a geodesic iff

$$(\nabla_{X^*} X^*)_{\tau_{\gamma(t)}(o)} = 0, \quad \text{for all } t \in \mathbb{R}.$$

Now, since $\tau_{\gamma(t)}$ is a diffeomorphism of G/H for every t and since ∇ is G-invariant, we have

$$(\tau_{\gamma(t)})_*(\nabla_{X^*} X^*) = \nabla_{(\tau_{\gamma(t)})_* X^*} (\tau_{\gamma(t)})_* X^*,$$

and from Proposition 1.9, we have

$$(\tau_{\gamma(t)})_* X^* = X^*,$$

so we obtain

$$(\tau_{\gamma(t)})_*(\nabla_{X^*} X^*) = \nabla_{X^*} X^*,$$

which evaluated at $\tau_{\gamma(t)}(o)$ yields

$$(\tau_{\gamma(t)})_*(\nabla_{X^*}X^*)_{\tau_{\gamma(t)(o)}} = (\nabla_{X^*}X^*)_{\tau_{\gamma(t)(o)}};$$

that is,

$$(d\tau_{\gamma(t)})_o(\nabla_{X^*}X^*)_o = (\nabla_{X^*}X^*)_{\tau_{\gamma(t)(o)}}.$$

Since $(d\tau_{\gamma(t)})_o$ is a bijection, we have $(\nabla_{X^*}X^*)_{\tau_{\gamma(t)}(o)} = 0$ for all $t \in \mathbb{R}$ iff $(\nabla_{X^*}X^*)_o = 0$ iff $\alpha(X, X) = 0$ for all $X \in \mathfrak{m}$, establishing our claim. $\square$

Since we showed that a G-invariant connection on G/H corresponds to a bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ whose skew-symmetric part α_S is given by

$$\alpha_S = \frac{1}{2}[X, Y]_{\mathfrak{m}},$$

if there is a G-invariant torsion-free connection on G/H such that the the curves $t \mapsto \tau_{e^{tX}}(o)$ are geodesics through o for all $X \in \mathfrak{m}$, then

$$\alpha(X, Y) = -\frac{1}{2}[X, Y]_{\mathfrak{m}}.$$

Conversely, because Ad_h is induced by the Lie group isomorphism $R_{h^{-1}} \circ L_h$, it is a Lie algebra isomorphism, so the Lie bracket $[X, Y]$ is Ad_h-invariant for all $h \in H$, and Theorem 1.8 shows that there is G-invariant connection induced by

$$\alpha(X, Y) = -\frac{1}{2}[X, Y]_{\mathfrak{m}}.$$

Now, if the curves $t \mapsto \tau_{e^{tX}}(o)$ are geodesics through o for all $X \in \mathfrak{m}$, since we have $d/dt(\tau_{e^{tX}(o)})|_{t=0} = X_o^*$, by the uniqueness of geodesics passing through o and with initial velocity X_o^*, we see that all geodesics through o are of the form $t \mapsto \tau_{e^{tX}}(o)$. Thus, we obtain the following result which is a version of Theorem 2.10 from Kobayashi and Nomizu [3] (Chapter X).

Theorem 1.11. *Given any reductive homogeneous manifold G/H with reductive decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, there is a unique G-invariant torsion-free connection ∇ on G/H such that all geodesics through o are given by the curves $t \mapsto \tau_{e^{tX}}(o)$ iff the bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ associated with ∇ is given by*

$$\alpha(X, Y) = -\frac{1}{2}[X, Y]_{\mathfrak{m}}, \quad \textit{for all } X, Y \in \mathfrak{m}.$$

We call the above connection the *Cartan connection* on G/H.

Remark: Theorem 2.10 In Kobayashi and Nomizu [3] states that

$$\alpha(X, Y) = \frac{1}{2}[X, Y]_{\mathfrak{m}},$$

with a + sign. This appears to be in contradiction with our result. The reason is that Kobayashi and Nomizu define the action vector field X^* associated with a vector $X \in \mathfrak{g}$ in terms of the *right* action of e^{tX} on G/H (see [2], page 42). We use the *left* action of e^{tX} on G/H (as most other authors of books written after the 1980's do).

The Levi–Civita connection is preserved by diffeomorphims, so in particular, any Levi–Civita connection on a homogeneous space is G-invariant. We also know that if G/H admits a G-invariant metric, then the Levi–Civita connection induced by that metric is given by

$$(d\pi_1)^{-1}(\nabla_{X^*}Y^*)_o = -\frac{1}{2}[X,Y]_{\mathfrak{m}} + U(X,Y),$$

where $[X,Y]_{\mathfrak{m}}$ is the component of $[X,Y]$ on $\mathfrak{m}$ and $U(X,Y)$ is determined by

$$2\langle U(X,Y), Z\rangle = \langle [Z,X]_{\mathfrak{m}}, Y\rangle + \langle X, [Z,Y]_{\mathfrak{m}}\rangle,$$

for all $Z \in \mathfrak{m}$. Therefore, we deduce that the Levi–Civita connection associated with a G-invariant metric on G/H coincides with the Cartan connection on G/H iff $U \equiv 0$ iff G/H is naturally reductive (see Kobayashi and Nomizu [3] (Chapter X, Theorem 3.3).

1.5 A Connection on $\mathbf{SE}(n)$

We compute the Levi-Civita connection associated with the left-invariant metric on $\mathbf{SE}(n)$ induced by the inner product in $\mathfrak{se}(n)$ given by

$$\langle X, Y\rangle = \mathrm{tr}(XY^\top) = \mathrm{tr}(X^\top Y).$$

For left-invariant vector fields, the inner products $\langle X, Y\rangle$ are constant, so the Koszul formula reduces to

$$2\langle \nabla_X Y, Z\rangle = \langle [X,Y], Z\rangle - \langle [Y,Z], X\rangle - \langle [X,Z], Y\rangle.$$

If

$$X = \begin{pmatrix} S_1 & u_1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} S_2 & u_2 \\ 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} S_3 & u_3 \\ 0 & 0 \end{pmatrix},$$

then we have

$$[Y,Z] = YZ - ZY = \begin{pmatrix} S_2S_3 & S_2u_3 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} S_3S_3 & S_3u_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} S_2S_3 - S_3S_2 & S_2u_3 - S_3u_2 \\ 0 & 0 \end{pmatrix},$$

and

$$\begin{aligned} \langle [Y,Z], X\rangle &= \mathrm{tr}\begin{pmatrix} S_2S_3 - S_3S_2 & S_2u_3 - S_3u_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} S_1^\top & 0 \\ u_1^\top & 0 \end{pmatrix} \\ &= \mathrm{tr}\begin{pmatrix} (S_2S_3 - S_3S_2)S_1^\top + (S_2u_3 - S_3u_2)u_1^\top & 0 \\ 0 & 0 \end{pmatrix} \\ &= \mathrm{tr}(S_2S_3S_1^\top - S_3S_2S_1^\top + S_2u_3u_1^\top - S_3u_2u_1^\top). \end{aligned}$$

Similarly,

$$\langle [X, Z], Y\rangle = \mathrm{tr}(S_1S_3S_2^\top - S_3S_1S_2^\top + S_1u_3u_2^\top - S_3u_1u_2^\top),$$

so we get

$$\begin{aligned}\langle [Y, Z], X\rangle + \langle [X, Z], Y\rangle = \mathrm{tr}(&S_2S_3S_1^\top - S_3S_2S_1^\top + S_1S_3S_2^\top - S_3S_1S_2^\top \\ &+ S_2u_3u_1^\top - S_3u_2u_1^\top + S_1u_3u_2^\top - S_3u_1u_2^\top)\end{aligned}$$

and since $S_1^\top = -S_1, S_2^\top = -S_2$, we obtain

$$\begin{aligned}\langle [Y, Z], X\rangle + \langle [X, Z], Y\rangle = \mathrm{tr}(&-S_2S_3S_1 + S_3S_2S_1 - S_1S_3S_2 + S_3S_1S_2 \\ &+ S_2u_3u_1^\top + S_1u_3u_2^\top - S_3(u_2u_1^\top + u_1u_2^\top)).\end{aligned}$$

Now, the first and the fourth terms cancel out since

$$\mathrm{tr}(S_2S_3S_1) = \mathrm{tr}(S_3S_1S_2),$$

and the second and the third terms cancel out since

$$\mathrm{tr}(S_3S_2S_1) = \mathrm{tr}(S_1S_3S_2).$$

Furthermore, because $u_2u_1^\top + u_1u_2^\top$ is symmetric and S_3 is skew symmetric, we have

$$\mathrm{tr}(S_3(u_2u_1^\top + u_1u_2^\top)) = 0.$$

Indeed, if S is a skew symmetric and H is a symmetric matrix

$$\mathrm{tr}(SH) = \mathrm{tr}((SH)^\top) = \mathrm{tr}(H^\top S^\top) = -\mathrm{tr}(HS) = -\mathrm{tr}(SH),$$

so $\mathrm{tr}(SH) = 0$. After simplifications, we get

$$\langle [Y, Z], X\rangle + \langle [X, Z], Y\rangle = \mathrm{tr}(S_2u_3u_1^\top + S_1u_3u_2^\top) = \mathrm{tr}(S_2^\top u_1u_3^\top + S_1^\top u_2u_3^\top).$$

Then, if we observe that

$$\mathrm{tr}(S_2^\top u_1u_3^\top + S_1^\top u_2u_3^\top) = \mathrm{tr}\begin{pmatrix} 0 & S_2^\top u_1 + S_1^\top u_2 \\ 0 & 0\end{pmatrix}\begin{pmatrix} S_3^\top & 0 \\ u_3^\top & 0\end{pmatrix},$$

we can write

$$\begin{aligned}2\langle \nabla_X Y, Z\rangle &= \langle [X, Y], Z\rangle - \langle [Y, Z], X\rangle - \langle [X, Z], Y\rangle \\ &= \langle [X, Y], Z\rangle - \left\langle \begin{pmatrix} 0 & S_2^\top u_1 + S_1^\top u_2 \\ 0 & 0\end{pmatrix}, Z\right\rangle \\ &= \langle [X, Y], Z\rangle + \left\langle \begin{pmatrix} 0 & S_2u_1 + S_1u_2 \\ 0 & 0\end{pmatrix}, Z\right\rangle,\end{aligned}$$

which yields

$$\nabla_X Y = \frac{1}{2}\left([X, Y] + \begin{pmatrix} 0 & S_2u_1 + S_1u_2 \\ 0 & 0 \end{pmatrix}\right).$$

Since

$$[X, Y] = \begin{pmatrix} S_1S_2 - S_2S_1 & S_1u_2 - S_2u_1 \\ 0 & 0 \end{pmatrix},$$

we also have

$$\nabla_X Y = \frac{1}{2}\begin{pmatrix} S_1S_2 - S_2S_1 & 2S_1u_2 \\ 0 & 0 \end{pmatrix}.$$

Consider the inner product

$$\langle X, Y \rangle = \mathrm{tr}(X^\top Y)$$

on $\mathfrak{se}(n)$. We claim that this inner product is invariant under the left action of $G = \mathbf{SE}(n)$. If

$$X = \begin{pmatrix} S & u \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad R = \begin{pmatrix} Q & z \\ 0 & 1 \end{pmatrix} \in \mathbf{SE}(n),$$

with $S^\top = -S$, $T^\top = -T$, $Q^\top Q = QQ^\top = I$, and $u, v, z \in \mathbb{R}^n$, then we have

$$\begin{pmatrix} Q & z \\ 0 & 1 \end{pmatrix}\begin{pmatrix} S & u \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} QS & Qu \\ 0 & 0 \end{pmatrix}$$
$$\begin{pmatrix} Q & z \\ 0 & 1 \end{pmatrix}\begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} QT & Qv \\ 0 & 0 \end{pmatrix},$$

so

$$\begin{aligned}
\langle RX, RY \rangle &= \mathrm{tr}\begin{pmatrix} S^\top Q^\top & 0 \\ u^\top Q^\top & 0 \end{pmatrix}\begin{pmatrix} QT & Qv \\ 0 & 0 \end{pmatrix} \\
&= \mathrm{tr}\begin{pmatrix} S^\top Q^\top QT & S^\top Q^\top Qv \\ u^\top Q^\top QT & u^\top Q^\top Qv \end{pmatrix} \\
&= \mathrm{tr}\begin{pmatrix} S^\top T & S^\top v \\ u^\top T & u^\top v \end{pmatrix} \\
&= \mathrm{tr}(S^\top T + u^\top v).
\end{aligned}$$

However

$$\langle X, Y \rangle = \mathrm{tr}\begin{pmatrix} S^\top & 0 \\ u^\top & 0 \end{pmatrix}\begin{pmatrix} T & v \\ 0 & 0 \end{pmatrix} = \mathrm{tr}\begin{pmatrix} S^\top T & S^\top v \\ u^\top T & u^\top v \end{pmatrix} = \mathrm{tr}(S^\top T + u^\top v),$$

which proves that

$$\langle RX, RY \rangle = \langle X, Y \rangle.$$

Acknowlegments:

Bibliography

[1] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry.* Universitext. Springer Verlag, second edition, 1993.

[2] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of Differential Geometry, I.* Wiley Classics. Wiley-Interscience, first edition, 1996.

[3] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of Differential Geometry, II.* Wiley Classics. Wiley-Interscience, first edition, 1996.

[4] Barrett O'Neill. *Semi-Riemannian Geometry With Applications to Relativity.* Pure and Applies Math., Vol 103. Academic Press, first edition, 1983.