

NOTES ON THE RIEMANNIAN GEOMETRY OF LIE GROUPS

Michael L. Geis ^a

VOLUME ,

Sponsored by

Rose-Hulman Institute of Technology

Department of Mathematics

Terre Haute, IN 47803

Email: mathjournal@rose-hulman.edu

<http://www.rose-hulman.edu/mathjournal>

^aRutgers University

ROSE-HULMAN UNDERGRADUATE MATHEMATICS JOURNAL

VOLUME ,

NOTES ON THE RIEMANNIAN GEOMETRY OF LIE GROUPS

Michael L. Geis

Abstract. Lie groups occupy a central position in modern differential geometry and physics, as they are very useful for describing the continuous symmetries of a space. This paper is an expository article meant to introduce the theory of Lie groups, as well as survey some results related to the Riemannian geometry of groups admitting invariant metrics. In particular, a non-standard proof of the classification of invariant metrics is presented. For those unfamiliar with tensor calculus, a section devoted to tensors on manifolds and the Lie derivative is included.

Acknowledgements: The author wishes to thank Dr. Po Lam Yung for his great encouragement and continued support.

1 Introduction and Notation

It is well known that many finite groups can be thought of quite explicitly as the symmetries of some geometric object. Lie groups extend this idea to groups whose elements are parametrized by several continuous (real) parameters $(x_1, \dots, x_n)$. For example, consider the group of invertible linear transformations from $\mathbb{R}^n$ to itself. This group can be identified with $\mathrm{GL}_n(\mathbb{R})$, the group of $n \times n$ invertible matrices, whose matrix elements can be parametrized by their n^2 real entries. If we take the more sophisticated viewpoint of symmetries as structure-preserving transformations, we can interpret $\mathrm{GL}_n(\mathbb{R})$ (as well as many other matrix groups) as symmetries of $\mathbb{R}^n$. For this, we have to have a specific structure in mind. To start, consider $\mathbb{R}^n$ only as an n -dimensional real vector space. $\mathrm{GL}_n(\mathbb{R})$ preserves addition, scalar multiplication, and the dimension of $\mathbb{R}^n$, so it preserves this structure. One can also view $\mathbb{R}^n$ as an inner product space, with the standard dot product. In this case, the subgroup $\mathrm{O}(n)$ of $\mathrm{GL}_n(\mathbb{R})$ consists of the linear transformations which preserve the vector space structure as well as this inner product structure.

Symmetries of this type have two useful properties. First, we can often exploit them to streamline computations. In the simple example of $\mathbb{R}^n$ as an inner product space, the $\mathrm{O}(n)$ symmetry tells us that the inner product *looks* the same with respect to any orthonormal basis, so we might as well choose the most convenient one to work with. This simple idea carries over to much more complicated situations where a clever choice of basis or coordinates can simplify things greatly. More importantly, symmetries often carry important intrinsic information. In dynamics, symmetries of a system are transformations which leave the Lagrangian invariant. (Take this to mean transformations which don't change the physics of the system). The famous Noether's theorem asserts, roughly speaking, that to each such symmetry there is a corresponding quantity which is constant as the system evolves in time. For example, suppose the Lagrangian does not depend on time. (This means the system behaves the same whether we set it in motion right now or next week). Then the Lagrangian is invariant under the Lie group of time translations $t \rightarrow t + t_0$. The conserved quantity corresponding to this symmetry is just the total energy of the system. Indeed, we learn in any basic physics class that the sum of the kinetic and the potential energy is a constant. The time-translation symmetry of the Lagrangian is the underlying reason.

This article is intended to be a presentation of a few results about the geometry of Lie groups which admit bi-invariant metrics, as well as a basic introduction to the general theory of Lie groups. Section 2 is devoted to developing some preliminary tools used to prove the results of interest, namely the Lie derivative and Killing vector fields. In section 3 we outline the basics of the theory of Lie groups, including the Lie algebra of a Lie group and the exponential map. We then focus on the special case of groups which admit bi-invariant metrics. We give a characterization of bi-invariant metrics which allows us to prove that all geodesics are translates of 1-parameter subgroups and as a consequence, that these groups are complete. Finally we use these results to prove that the sectional curvature of such

groups is non-negative.

1.1 Notation

Throughout, M will always denote a topological manifold of dimension n with a smooth (C^∞) structure. Recall that such an M is locally homeomorphic to $\mathbb{R}^n$. By a coordinate chart at $p \in M$, we mean a particular choice of an open set $U \subset \mathbb{R}^n$ and a homeomorphism $\mathbf{x}$ mapping U onto an open neighborhood of p . Such a choice gives coordinates on $\mathbf{x}(U)$ by identifying $(x_1, \dots, x_n) \in U$ with $\mathbf{x}(x_1, \dots, x_n)$. For brevity, we sometimes use the term *local coordinates* $(x_1, \dots, x_n)$ at p to mean a coordinate chart at p , suppressing the actual homeomorphism. For $p \in M$, we denote the tangent space of M at p by $T_p M$. A function $f: M \rightarrow \mathbb{R}$ is smooth if for any coordinate chart $\mathbf{x}: U \subset \mathbb{R}^n \rightarrow M$ the function $f \circ \mathbf{x}$ is smooth on U . Smooth functions on M form a ring which we denote by $C^\infty(M)$. If N is another manifold and $\varphi: M \rightarrow N$ is a smooth map, $d\varphi_p: T_p M \rightarrow T_{\varphi(p)} N$ is its differential at p . If $I \subset \mathbb{R}$ and $\alpha: I \rightarrow M$ is a smooth curve then we write $\alpha'(t)$ to mean $d\alpha_t\left(\frac{\partial}{\partial t}\right)$, where $\frac{\partial}{\partial t} \in T_t I$.

A smooth vector field X on M is an assignment to every $p \in M$ a vector $X(p) \in T_p M$ so that given any local coordinates $\mathbf{x}: U \subset \mathbb{R}^n \rightarrow M$, if $X = \sum a_i \frac{\partial}{\partial x^i}$ on $\mathbf{x}(U)$, the component functions a_i are smooth on $\mathbf{x}(U)$. A smooth vector field X acts like a linear differential operator on smooth functions, assigning to each $f \in C^\infty(M)$ a new function $X(f)$ which, in local coordinates is equal to

$$X(f) = \sum a_i \frac{\partial \tilde{f}}{\partial x^i} \quad (1)$$

where $\tilde{f} = f \circ \mathbf{x}$ is the expression of f in these coordinates. If X and Y are smooth vector fields on M , their commutator bracket is defined to be $[X, Y] = XY - YX$ and is another smooth vector field. We will denote the space of smooth vector fields on M by $\mathcal{X}(M)$. Throughout, we will drop the adjective smooth and assume all vector fields are smooth unless stated otherwise.

A manifold M is Riemannian if each tangent space has an inner product which varies smoothly across M (in the sense of tensors, see definition 2.2). We will refer to the metric on M by either $\langle \cdot, \cdot \rangle$ or $g(\cdot, \cdot)$. Finally, recall that the Levi-Civita connection on M is the mapping

$$\begin{aligned} \nabla: \mathcal{X}(M) \times \mathcal{X}(M) &\rightarrow \mathcal{X}(M) \\ (X, Y) &\rightarrow \nabla_X Y \end{aligned}$$

determined uniquely by the following properties holding for all $f, g \in C^\infty(M)$ and $X, Y, Z \in \mathcal{X}(M)$

$$(i) \quad \nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$$

- (ii) $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$
- (iii) $\nabla_X (fY) = X(f)Y + f\nabla_X Y$
- (iv) $\nabla_X Y - \nabla_Y X = [X, Y]$
- (v) $Z(g(X, Y)) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$

It can be shown by a computation in local coordinates that for vector fields X and Y on M , the value of $\nabla_X Y$ at p depends only on $X(p)$ and the value of Y in a neighborhood of p . Thus we can speak of $\nabla_X Y$ when X and Y are vector fields defined on some open set. More generally If $\alpha : I \subset \mathbb{R} \rightarrow M$ is a curve and $V(t)$ is a vector field along $\alpha(t)$, there is a unique notion of the time derivative of V , written $\frac{DV}{dt}$, such that whenever V is the restriction of a vector field defined in a neighborhood of $\alpha(I)$,

$$\frac{DV}{dt}(t) = \nabla_{\alpha'(t)} V$$

α is a *geodesic* if $\frac{D\alpha'}{dt} = 0$. That is, if its tangent vector field is *constant* along it, with respect to this new notion of derivative. (See [3])

2 Tensors on Manifolds

Definition 2.1. Let V be a finite dimensional real vector space and let r and s be non-negative integers. An (r, s) tensor on V is a multilinear mapping $A : V \times \cdots \times V \times V^* \times \cdots \times V^* \rightarrow \mathbb{R}$ (with V appearing r times and V^* appearing s times)

The numbers r and s are called the covariant and contravariant orders of A . We say that a tensor is covariant if its contravariant order is zero and contravariant if its covariant order is zero. The set of all (r, s) tensors on V forms a vector space under pointwise addition and scalar multiplication of maps which we will denote by $\mathcal{T}_s^r(V)$.

Proposition 2.1. Let V be a finite dimensional real vector space and let $\{e_1, \dots, e_n\}$ be a basis for V with dual basis $\{\epsilon^1, \dots, \epsilon^n\}$. The tensors $\{\epsilon^{i_1} \otimes \cdots \otimes \epsilon^{i_r} \otimes e_{j_1} \otimes \cdots \otimes e_{j_s}\}$ (where each i_k and j_l range from 1 to n) form a basis for $\mathcal{T}_s^r(V)$.

This means that every (r, s) tensor A on V can be written uniquely as

$$A = \sum A_{i_1 \dots i_r}^{j_1 \dots j_s} \epsilon^{i_1} \otimes \cdots \otimes \epsilon^{i_r} \otimes e_{j_1} \otimes \cdots \otimes e_{j_s}$$

Where the sum is taken over all $i_1, \dots, i_r$ and $j_1, \dots, j_s$ from 1 to n . The coefficients $A_{i_1 \dots i_r}^{j_1 \dots j_s}$ are just the values of A on those basis vectors with matching indices. That is,

$$A_{i_1 \dots i_r}^{j_1 \dots j_s} = A(e_{i_1}, \dots, e_{i_r}, \epsilon^{j_1}, \dots, \epsilon^{j_s})$$

Definition 2.2. Let M be a smooth manifold. A smooth (r, s) tensor field A on M is an assignment to each $p \in M$ an (r, s) tensor on $T_p M$, $p \mapsto A(p)$ so that given any local coordinates $\mathbf{x}: U \subset \mathbb{R}^n \rightarrow M$, if

$$A = \sum A_{i_1 \dots i_r}^{j_1 \dots j_s} dx^{i_1} \otimes \dots \otimes dx^{i_r} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{j_s}}$$

on $\mathbf{x}(U)$ then the component functions $A_{i_1 \dots i_r}^{j_1 \dots j_s}$ are smooth on $\mathbf{x}(U)$.

Example. Let g be a Riemannian metric on M . Then g is a smooth (positive definite and symmetric) $(2, 0)$ tensor field on M , sometimes called the metric tensor. It is an assignment to each point $p \in M$ an inner product $\langle \cdot, \cdot \rangle_p$ on $T_p M$ such that for any coordinate chart $\mathbf{x}: U \subset \mathbb{R}^n \rightarrow M$, putting $q = \mathbf{x}(x_1, \dots, x_n)$, the functions

$$g_{ij}(x_1, \dots, x_n) = \left\langle \frac{\partial}{\partial x^i}(q), \frac{\partial}{\partial x^j}(q) \right\rangle_q$$

are smooth on U for every choice of i and j . The g_{ij} are exactly the functions described in the previous definition since we can write $g = \sum g_{ij} dx^i \otimes dx^j$.

There is an equivalent way to view a tensor field on a manifold which very closely resembles a tensor on a vector space. For notational simplicity, we restrict to the case of covariant tensor fields. Let A be a smooth $(r, 0)$ tensor field on a manifold M . A associates to every r -tuple of smooth vector fields $X_1, \dots, X_r$ on M a function $A(X_1, \dots, X_r)$ defined by

$$A(X_1, \dots, X_r)(p) = A(p)(X_1(p), \dots, X_r(p)) \quad (2)$$

The key property of this association is that it is actually multi-linear with respect to functions on M , meaning that

$$A(X_1, \dots, fX_i + gY_i, \dots, X_r) = fA(X_1, \dots, X_i, \dots, X_r) + gA(X_1, \dots, Y_i, \dots, X_r)$$

for any functions f and g . Using this multi-linearity property, one can check that in coordinates

$$A(X_1, \dots, X_r) = \sum A_{i_1 \dots i_r} x_1^{i_1} \dots x_r^{i_r}$$

where $x_j^{i_j}$ is the i_j^{th} coefficient of X_j . Hence the function $A(X_1, \dots, X_r)$ is smooth. Therefore every smooth covariant tensor field A induces a mapping $A: \mathcal{X}(M) \times \dots \times \mathcal{X}(M) \rightarrow C^\infty(M)$ which is multi-linear with respect to smooth functions. Actually, it is true that such a mapping is multi-linear with respect to smooth functions if and only if it is induced by some smooth tensor field. Thus one could define a tensor on a manifold M as in definition 2.1 changing V to $\mathcal{X}(M)$ and $\mathbb{R}$ to the ring $C^\infty(M)$.

2.1 The Lie Derivative

There are many important objects in differential geometry which are tensors, the Riemannian metric and Riemann curvature tensor being two notable examples. It turns out that that one would often like to describe how a tensor changes as one moves along some fixed direction on a manifold, which leads to problem of defining a notion of directional derivative for tensors.

Let us first consider the simplest case, that of a $(0,1)$ tensor, or a vector field. As a motivating example, consider the directional derivative of a vector field X on $\mathbb{R}^n$ at the point p , with respect to a fixed single vector $v \in \mathbb{R}^n$. Thinking of such a vector field as a mapping $X: \mathbb{R}^n \rightarrow \mathbb{R}^n$, we might define

$$D_v X = \lim_{t \rightarrow 0} \frac{X(p + tv) - X(p)}{t} \quad (3)$$

It is immediate that this definition is extremely problematic, because it depends almost entirely upon the fact that $\mathbb{R}^n$ is a vector space. Specifically, we are using the fact that each tangent space of $\mathbb{R}^n$ can be identified with $\mathbb{R}^n$ itself by an affine translation. ($T_p \mathbb{R}^n$ can be thought of as all of the vectors of $\mathbb{R}^n$ beginning at p instead of the origin.) Indeed, the vectors $X(p + tv)$ and $X(p)$ are not even in the same tangent space, and it is only this identification that allows us to subtract them! If we are to generalize this definition, we will need some way of connecting two different tangent spaces on arbitrary manifolds, without utilizing any vector space structure.

Theorem 2.1. (*Fundamental Theorem of Ordinary Differential Equations*) *Let M be a smooth manifold and let X be a vector field on M . For each $p \in M$, there is a neighborhood $U \subset M$ of p , a real number $\delta > 0$, and a smooth mapping $\phi: (-\delta, \delta) \times U \rightarrow M$ such that for each $q \in U$, $t \mapsto \phi(t, q)$ is the unique curve satisfying $\phi'(t, q) = X(\phi(t, q))$ and $\phi(0, q) = q$*

For each $t \in (-\delta, \delta)$ the family of mappings $\phi_t: U \rightarrow M$ given by $\phi_t(q) = \phi(t, q)$ is called the *local flow* of X . Geometrically, ϕ_t sends each $q \in U$ to the point obtained by moving along the integral curve of X through q for a time t . The local flow is a diffeomorphism onto $\phi_t(U)$ with inverse ϕ_{-t} . We will see more about flows of vector fields later on, but for now, the importance of the local flow is that it can be used to generalize our definition of directional derivative to arbitrary manifolds. For more on theorem 2.1 see [3].

Definition 2.3. *Let M be a smooth manifold and X and Y be vector fields on M . For each $p \in M$, let $U \subset M$ be a neighborhood of p where the local flow of X , x_t is defined for $t \in (-\delta, \delta)$. Define the Lie derivative of Y with respect to X to be the vector field $\mathcal{L}_X Y$ given by*

$$(\mathcal{L}_X Y)(p) = \lim_{t \rightarrow 0} \frac{dx_{-t}(Y(x_t(p))) - Y(p)}{t} \quad (4)$$

Observe that this definition makes sense, since for each t such that x_t is defined, dx_{-t} pulls back the vector $Y(x_t(p))$ to a vector in $T_p M$. This means that the subtraction defining $(\mathcal{L}_X Y)(p)$ makes sense, and that it is a vector in $T_p M$. It is not hard to show that $\mathcal{L}_X Y$ is actually a smooth vector field. (Intuitively, this is because the local flow depends smoothly on the base point p). Also, notice that $(\mathcal{L}_X Y)(p)$ is measuring the amount Y changes as one moves an infinitesimal distance along the integral curve of X at p , that is, as one moves an infinitesimal distance in the direction of $X(p)$. Although this definition looks quite intractable, there is an alternative computationally simple expression of the Lie derivative.

Proposition 2.2. *Let X and Y be smooth vector fields on a manifold M . Then $\mathcal{L}_X Y = [X, Y]$.*

See [3] for a proof of this proposition and more about the commutator bracket. At this point we have succeeded in generalizing the directional derivative in $\mathbb{R}^n$ to vector fields on manifolds. Indeed, $D_v X$ in our motivating example (3) is simply $\mathcal{L}_V X$, where V is the constant vector field taking the value v everywhere. But we wish to further generalize this derivative to arbitrary tensors. Let $U \subset M$ be an open subset of M and $\phi: U \rightarrow \phi(U)$ be a diffeomorphism. If X is a smooth vector field defined on $\phi(U)$, recall that the pullback of X by ϕ , $\phi^* X$, is the vector field on U defined by

$$(\phi^* X)(p) = d\phi_{\phi(p)}^{-1}(X(\phi(p)))$$

Using this language, the definition of the Lie derivative is simply

$$\mathcal{L}_X Y(p) = \lim_{t \rightarrow 0} \frac{(x_t^* Y)(p) - Y(p)}{t} = \left. \frac{d}{dt} ((x_t^* Y)(p)) \right|_{t=0} \quad (5)$$

Since we are interested in extending the Lie derivative to general tensors, we might try defining the pullback of a tensor by a diffeomorphism and take (5) as the definition. Again we shall only consider covariant tensors for notational convenience.

Definition 2.4. *Let U be an open subset of M and $\phi: U \rightarrow M$ be a smooth map. If A is an $(r, 0)$ tensor field defined on a neighborhood of $\phi(U)$, then the pullback of A by ϕ , $\phi^* A$, is the $(r, 0)$ tensor field on U defined by*

$$(\phi^* A)(p)(v_1, \dots, v_r) = A(\phi(p))(d\phi_p(v_1), \dots, d\phi_p(v_r))$$

for each $p \in U$ and $v_1, \dots, v_r \in T_p M$

Remark. *Covariant tensors can be pulled back by any smooth map, however mixed (r, s) tensors can only be pulled back by diffeomorphisms.*

Definition 2.5. (The Lie Derivative) Let M be a smooth manifold, X be a vector field on M , and A be an $(r, 0)$ tensor field on M . For each fixed $p \in M$, let $U \subset M$ be a neighborhood of p where the local flow of X , x_t , is defined for $t \in (-\delta, \delta)$. Define the Lie derivative of A with respect to X to be the $(r, 0)$ tensor field $\mathcal{L}_X A$ defined by

$$(\mathcal{L}_X A)(p)(v_1, \dots, v_r) = \left. \frac{d}{dt} ((x_t^* A)(p)(v_1, \dots, v_r)) \right|_{t=0} \quad (6)$$

for all $v_1, \dots, v_r \in T_p M$

Proposition 2.3. Let M be a smooth manifold, X be a vector field on M and $f \in C^\infty(M)$. Then $\mathcal{L}_X f = X(f)$.

Proof. Fix $p \in M$. Let $U \subset M$ be a neighborhood of p such that the local flow of X , x_t is defined on U for $t \in (-\delta, \delta)$. By definition,

$$(\mathcal{L}_X f)(p) = \left. \frac{d}{dt} ((x_t^* f)(p)) \right|_{t=0}$$

and $x_t^* f$ is the function $f \circ x_t$, so

$$(\mathcal{L}_X f)(p) = \left. \frac{d}{dt} (f \circ x_t(p)) \right|_{t=0}$$

Writing f and the curve $t \mapsto x_t(p)$ in local coordinates and differentiating the above expression, one gets exactly the coordinate expression (1) for $X(f)(p)$. □

The next proposition is a computationally useful one which relates the Lie derivative of a covariant tensor to the Lie derivative of its arguments.

Proposition 2.4. Let M be a smooth manifold, A be an $(r, 0)$ tensor on M , and $X_1, \dots, X_r$, and Y be vector fields on M . Then

$$\mathcal{L}_Y (A(X_1, \dots, X_r)) = (\mathcal{L}_Y A)(X_1, \dots, X_r) + \sum_{i=1}^r A(X_1, \dots, \mathcal{L}_Y X_i, \dots, X_r)$$

Proof. We verify the equality pointwise. Fix $p \in M$ and choose a neighborhood $U \subset M$ of p and local coordinates $(u_1, \dots, u_n)$ inside U such that: the local flow of Y , y_t is defined on U for all $t \in (-\delta, \delta)$, $Y = \frac{\partial}{\partial u^i}$, $A = \sum A_{i_1, \dots, i_r} du^{i_1} \otimes \dots \otimes du^{i_r}$, and $X_i = \sum x_i^j \frac{\partial}{\partial u^j}$ in U . Note by our choice of coordinates, $y_t(u_1, \dots, u_n) = (u_1 + t, \dots, u_n)$. Therefore the matrix of $d(y_t)_q$ relative to these coordinates is the identity for all $q \in U$.

Consider the left hand side of the equality. By proposition 2.4, $\mathcal{L}_Y (A(X_1, \dots, X_r))(p)$ is equal to

$$\begin{aligned}
Y(A(X_1, \dots, X_r))(p) &= \sum \frac{\partial}{\partial u^1} (A_{i_1, \dots, i_r} x_1^{i_1} \cdots x_r^{i_r}) \Big|_p \\
&= \sum \frac{\partial A_{i_1, \dots, i_r}}{\partial u^1} x_1^{i_1} \cdots x_r^{i_r} \Big|_p + \sum_{j=1}^r \left(\sum A_{i_1, \dots, i_r} x_1^{i_1} \cdots \frac{\partial x_j^{i_r}}{\partial u^1} \cdots x_r^{i_r} \Big|_p \right)
\end{aligned}$$

Now for the first term on the right hand side, $(\mathcal{L}_Y A)(X_1, \dots, X_r)(p) = (\mathcal{L}_Y A)(p)(X_1(p), \dots, X_r(p))$ and by definition

$$(\mathcal{L}_Y A)(p)(X_1(p), \dots, X_r(p)) = \frac{d}{dt} \left(A(y_t(p)) \left(d(y_t)_p(X_1(p)), \dots, d(y_t)_p(X_r(p)) \right) \right) \Big|_{t=0}$$

Since the matrix of $d(y_t)_p$ is the identity in these coordinates

$$d(y_t)_p(X_i(p)) = \sum_j x_i^j(p) \frac{\partial}{\partial u^j}(y_t(p))$$

So the last expression becomes

$$\frac{d}{dt} \left(\sum A_{i_1, \dots, i_r}(y_t(p)) x_1^{i_1}(p) \cdots x_r^{i_r}(p) \right) \Big|_{t=0}$$

If p corresponds to $(u_1^0, \dots, u_n^0)$, then $y_t(p) = (u_1^0 + t, \dots, u_n^0)$. Inserting this expression, differentiating, and setting $t = 0$, we get

$$(\mathcal{L}_Y A)(X_1, \dots, X_r)(p) = \sum \frac{\partial A_{i_1, \dots, i_r}}{\partial u^1} x_1^{i_1} \cdots x_r^{i_r} \Big|_p$$

For the final term, note that by proposition 2.3

$$\mathcal{L}_Y X_i = \left[\frac{\partial}{\partial u^1}, X_i \right] = \frac{\partial}{\partial u^1} X_i - X_i \frac{\partial}{\partial u^1}$$

Or, in coordinates

$$\begin{aligned}
\mathcal{L}_Y X_i &= \sum_j \left(\frac{\partial x_i^j}{\partial u^1} \frac{\partial}{\partial u^j} + x_i^j \frac{\partial}{\partial u^1 \partial u^j} \right) - x_i^j \frac{\partial}{\partial u^j \partial u^1} \\
&= \sum_j \frac{\partial x_i^j}{\partial u^1} \frac{\partial}{\partial u^j}
\end{aligned}$$

by equality of mixed partials. Therefore

$$\sum_{j=1}^r A(X_1, \dots, \mathcal{L}_Y X_j, \dots, X_r)(p) = \sum_{j=1}^r \left(\sum A_{i_1, \dots, i_r} x_1^{i_1} \cdots \frac{\partial x_j^{i_j}}{\partial u^1} \cdots x_r^{i_r} \Big|_p \right)$$

The expressions for both terms on the right hand side together are exactly the expression obtained for the left hand side. □

Before moving on to the study of Lie groups, we make a short note about a special class of vector fields on Riemannian manifolds which we will need later on.

Definition 2.6. *Let M be a Riemannian manifold. A vector field X on M is a Killing field if for each $p \in M$, the local flow of X at p , $x_t: U \rightarrow M$, is an isometry for all t such that it is defined. That is, if*

$$\langle u, v \rangle_q = \langle d(x_t)_q(u), d(x_t)_q(v) \rangle_{x_t(q)}$$

for all $q \in U$, $u, v \in T_p M$

Proposition 2.5. *Let M be a Riemannian manifold and let g be the metric tensor on M . A vector field X is a Killing field if and only if $(\mathcal{L}_X g)(p)$ is the zero tensor for every $p \in M$.*

Proof. First, suppose that X is a Killing field. Fix $p \in M$ and let x_t be the local flow of X in the neighborhood U of p , for $t \in (-\delta, \delta)$. For any $u, v \in T_p M$, we have

$$\begin{aligned} (\mathcal{L}_X g)(p)(u, v) &= \frac{d}{dt} ((x_t^* g)(p)(u, v)) \Big|_{t=0} \\ &= \frac{d}{dt} \left(g(x_t(p)) \left(d(x_t)_p(u), d(x_t)_p(v) \right) \right) \Big|_{t=0} \\ &= \frac{d}{dt} (g(p)(u, v)) \Big|_{t=0} \\ &= 0 \end{aligned}$$

Thus, $(\mathcal{L}_X g)(p)$ is the zero tensor at every $p \in M$. For the converse, suppose that $(\mathcal{L}_X g)(p)$ is the zero tensor for every $p \in M$. Fix p and let x_t be the local flow of X on a neighborhood U of p for $t \in (-\delta, \delta)$. Let $t_0 \in (-\delta, \delta)$ and $u, v \in T_p M$ be arbitrary. Put $u' = d(x_{t_0})_p(u)$ and $v' = d(x_{t_0})_p(v)$. Now by assumption we have $(\mathcal{L}_X g)(x_{t_0}(p))(u', v') = 0$, and by definition

$$\begin{aligned} (\mathcal{L}_X g)(x_{t_0}(p))(u', v') &= \frac{d}{dt} ((x_t^* g)(x_{t_0}(p))(u', v')) \Big|_{t=0} \\ &= \frac{d}{dt} \left(g(x_t(x_{t_0}(p))) \left(d(x_t)_{x_{t_0}(p)} u', d(x_t)_{x_{t_0}(p)} v' \right) \right) \Big|_{t=0} \\ &= \frac{d}{dt} \left(g(x_{t+t_0}(p)) \left(d(x_{t+t_0})_p(u), d(x_{t+t_0})_p(v) \right) \right) \Big|_{t=0} \end{aligned}$$

Putting $s = t + t_0$, the last line becomes

$$\left. \frac{d}{ds} \left(g(x_s(p)) \left(d(x_s)_p(u), d(x_s)_p(v) \right) \right) \right|_{s=t_0}$$

So that this derivative vanishes for all $u, v \in T_p M$ and all $t_0 \in (-\delta, \delta)$, hence $g(x_s(p)) \left(d(x_s)_p(u), d(x_s)_p(v) \right)$ is constant on that interval. In particular, it is equal to its value at $s = 0$; but this says exactly that the local flow is an isometry at p for all $t \in (-\delta, \delta)$. Thus X is a Killing field. $\square$

3 Lie Groups

Simply put, a Lie group is both a group and a smooth manifold where the group and differentiable structures are compatible. Lie groups and Lie algebras are extremely important objects in differential geometry and modern physics, with their widespread utility due to the fact that they encode continuous symmetries of geometric spaces. Indeed, one often seeks to understand the geometry of a certain manifold by understanding its symmetries, in which actions of Lie groups and their linearizations (Lie algebras) are indispensable tools.

Definition 3.1. *A Lie group G is a group and a smooth manifold such that group multiplication $G \times G \rightarrow G$ $(x, y) \mapsto xy$ and group inversion $G \rightarrow G$ $x \mapsto x^{-1}$ are smooth maps.*

If G and H are Lie groups, a Lie group homomorphism $\varphi: G \rightarrow H$ is a smooth mapping which is also a homomorphism of the abstract groups. If the mapping is a diffeomorphism, then φ is called an isomorphism. Much of the structure of Lie groups comes from the so-called left and right translations.

Definition 3.2. *Let G be a Lie group, $s \in G$. The left translation by s is the map $L_s: G \rightarrow G$ is given by $L_s(t) = st$ for every $t \in G$. Right translations are defined analogously.*

The group structure implies that for every s , L_s and R_s are bijections with inverses $L_{s^{-1}}$ and $R_{s^{-1}}$. The conditions in definition 3.1 imply that both of these maps (and their inverses) are smooth. Thus, left and right translations are diffeomorphisms of G onto itself.

Familiar examples of Lie groups include $\mathbb{R}^n$ under addition and the elementary matrix groups ($\mathrm{GL}(n, \mathbb{R})$, $\mathrm{SL}(n, \mathbb{R})$, $\mathrm{O}(n)$, $\mathrm{U}(n)$, etc.). In fact, all of the matrix groups can be viewed as closed Lie subgroups of $\mathrm{GL}(n, \mathbb{C})$. In particular, the circle group $\mathbb{S}^1 \cong \mathrm{U}(1)$ is a Lie group. The product of Lie groups given the standard differentiable and group structures is again a Lie group, so the k -torus when viewed as the k -fold product of $\mathbb{S}^1$ is another example.

3.1 The Lie Algebra of a Lie Group

The study of Lie algebras is a very rich subject in its own right, but also an extremely powerful tool for understanding the geometry of Lie groups. Every Lie group has a Lie algebra which is intimately associated with it; many times information about the Lie algebra of a group directly translates into information about the group itself. Indeed, a common strategy involves translating a problem about Lie groups into a problem about their Lie algebras, which has the advantage of reducing an algebrogeometric situation down to a purely algebraic one. Many theorems exist which illustrate this powerful link between Lie groups and Lie algebras. For instance, the Lie algebra of each connected subgroup of a Lie group G is a subalgebra of the Lie algebra of G .

Definition 3.3. A (real) Lie algebra $\mathfrak{a}$ is a vector space over $\mathbb{R}$ equipped with a bilinear operation $[\cdot, \cdot]: \mathfrak{a} \times \mathfrak{a} \rightarrow \mathfrak{a}$ called the bracket, which satisfies the following properties:

$$(i) \quad [x, y] = -[y, x] \quad \forall x, y \in \mathfrak{a}$$

$$(ii) \quad [x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0 \quad \forall x, y, z \in \mathfrak{a}$$

Property (ii) is often referred to as the Jacobi identity. A subalgebra of a Lie algebra $\mathfrak{a}$ is a subspace of $\mathfrak{a}$ which is closed under the bracket operation. Two Lie algebras $\mathfrak{a}$ and $\mathfrak{b}$ are isomorphic when there exists a linear isomorphism between them which preserves the bracket. For us, the important thing about Lie algebras is that every Lie group G has an associated Lie algebra. Actually, every smooth manifold M has an associated Lie algebra, the Lie algebra of smooth vector fields $\mathcal{X}(M)$ equipped with the commutator bracket. However, the key point is that $\mathcal{X}(M)$ is infinite dimensional, while a Lie group has a distinguished finite dimensional Lie algebra. To see this, we must examine a special class of vector fields on G .

Let X be a vector field on G . Since the left translations are diffeomorphisms of G , for each $s \in G$, the pullback of X by L_s , L_s^*X , is defined.

Definition 3.4. A vector field X on a Lie group G , not necessarily smooth, is left invariant if for every $s \in G$, $L_s^*X = X$. That is, the pullback of X by every left translation coincides with X .

Stated differently, if X is left invariant, then $d(L_s)_t(X(t)) = X(st)$ for every $s, t \in G$. In particular, for $t = e$ we have that

$$d(L_s)_e(X(e)) = X(s) \tag{7}$$

for each $s \in G$. Hence a left invariant vector field is completely determined by its value at e .

Proposition 3.1. Let G be a Lie group and $\mathfrak{g}$ be the set of left invariant vector fields on G .

- (i) *Left invariant vector fields are smooth*
- (ii) *$\mathfrak{g}$ forms a vector space over $\mathbb{R}$ which is isomorphic to $T_e G$*
- (iii) *The bracket of two left invariant fields is left invariant*
- (iv) *$\mathfrak{g}$ is a Lie algebra with the the commutator bracket.*

Proof. (i) Left invariant vector fields are smooth

Fix $s \in G$. We verify that when X is written in local coordinates at s , then its coefficients are smooth functions of these coordinates. Choose local coordinates $\mathbf{x}: U \subset \mathbb{R}^n \rightarrow G$ at s and $\mathbf{y}: V \subset \mathbb{R}^n \rightarrow G$ at e . Then we also have local coordinates $\mathbf{z}: U \times V \rightarrow G \times G$ at (s, e) . Since group multiplication is smooth, its expression in these local coordinates $\varphi: U \times V \rightarrow U$,

$$\varphi(x_1, \dots, x_n, y_1, \dots, y_n) = (u_1, \dots, u_n)$$

is a smooth map. Now since X is left invariant, $X(t) = d(L_t)_e(X(e))$, and for $t \in \mathbf{x}(U)$, writing $\mathbf{x}^{-1}(t) = (x_1, \dots, x_n)$, the matrix of $d(L_t)_e$ relative to the coordinate basis $\{\frac{\partial}{\partial y^i}\}$ is

$$(d(L_t)_e)_{ij} = \left(\frac{\partial u_i}{\partial y^j} (x_1, \dots, x_n, y_1^0, \dots, y_n^0) \right)$$

where $\mathbf{y}^{-1}(e) = (y_1^0, \dots, y_n^0)$. Therefore, if $X(e) = \sum_i x_0^i \frac{\partial}{\partial y^i}$ then

$$X(t) = \sum_k \left(\sum_j \frac{\partial u_k}{\partial y^j} (x_1, \dots, x_n, y_1^0, \dots, y_n^0) x_0^j \right) \frac{\partial}{\partial x^k}$$

Since the u_k are smooth functions of $x_1, \dots, x_n$, so are its derivatives.

- (ii) $\mathfrak{g}$ inherits a vector space structure from the space of all vector fields on G . To see that it is isomorphic to $T_e G$, consider the mapping $\alpha: \mathfrak{g} \rightarrow T_e G$ given by $\alpha(X) = X(e)$. Clearly α is linear, and it is injective, since left invariant fields are uniquely determined by their values at e . It is also surjective; given $v \in T_e G$, define a vector field X by $X(s) = d(L_s)_e v$. Then X is left invariant, as one can check, and $X(e) = v$.
- (iii) Let $X, Y \in \mathfrak{g}$. We must show that $d(L_s)_t([X, Y](t)) = [X, Y](st)$ for all $s, t \in G$. To do this, we verify that they coincide as operators on functions. For arbitrary $s, t \in G$, and $f \in C^\infty(G)$:

$$\begin{aligned}
d(L_s)_t [X, Y](t)(f) &= [X, Y](t)(f \circ L_s) \\
&= X(t)(Y(f \circ L_s)) - Y(t)(X(f \circ L_s)) \\
&= X(t)(Y \circ L_s)(f) - Y(t)(X \circ L_s)(f) \\
&= X(t)(Y(f) \circ L_s) - Y(t)(X(f) \circ L_s) \\
&= (d(L_s)_t X(t))(Y(f)) - (d(L_s)_t Y(t))(X(f)) \\
&= X(st)(Y(f)) - Y(st)(X(f)) \\
&= [X, Y](st)(f)
\end{aligned}$$

Therefore $d(L_s)_t [X, Y](t) = [X, Y](st)$.

- (iv) By (iii), the Lie bracket when restricted to left invariant vector fields satisfies all of the properties of the bracket of a Lie algebra, since it is anti-commuting and satisfies the Jacobi identity. Hence $\mathfrak{g}$ is a real Lie algebra of the same dimension as G . □

Occasionally it is useful to view the Lie algebra of G as $T_e G$. To make the isomorphism in (ii) into a Lie algebra isomorphism, define the bracket of two vectors $u, v \in T_e G$ by $[u, v] = [U, V](e)$, where U and V are the left invariant fields corresponding to u, v .

Example. The Lie algebra of $\mathbb{R}$

The real numbers form a Lie group under addition with their standard differentiable structure. Since $\mathbb{R}$ is abelian, left and right translations coincide; they are the maps $L_s(t) = R_s(t) = t + s$. The tangent space of $\mathbb{R}$ at zero is the set $T_0 \mathbb{R} = \{\lambda \frac{\partial}{\partial t}(0) : \lambda \in \mathbb{R}\}$. Using the correspondence (7), and noting that for every t , $d(L_s)_t$ is clearly the identity matrix in coordinates, the left invariant fields on $\mathbb{R}$ are exactly the constant fields, $X(t) = \lambda \frac{\partial}{\partial t}(t)$ for some $\lambda \in \mathbb{R}$.

3.2 Flows of Left Invariant Vector Fields, 1-Parameter Subgroups

The aim of this section is to develop the most important link between a Lie group and its Lie algebra, the exponential map.

Let X be a vector field on a smooth manifold M and let p be a point of M . Recall that an integral curve of X through the point p is a smooth curve $\alpha : I \subset \mathbb{R} \rightarrow M$ such that $\alpha(0) = p$ and $\alpha'(t) = X(\alpha(t))$ for all $t \in I$. For any $p \in M$, theorem 2.1 guarantees the existence (and uniqueness) of such a curve on $(-\epsilon, \epsilon)$ for some $\epsilon > 0$. However, an integral curve may cease to exist outside some finite interval of time. The vector field $X(t) = t^2 \frac{\partial}{\partial t}$ on $\mathbb{R}$ is a simple example.

Definition 3.5. A vector field X on M is called complete if every integral curve of X is defined for all $t \in \mathbb{R}$.

Proposition 3.2. Every left invariant vector field on a Lie group G is complete.

Proof. Let $X \in \mathfrak{g}$ and let γ be the integral curve of X through e (Theorem 2.1 guarantees it is defined on $(-\epsilon, \epsilon)$ for some $\epsilon > 0$). For any $s \in G$, the integral curve of X through s is $\gamma_s(t) = s \cdot \gamma(t)$. This follows from the uniqueness of integral curves, the fact that $\gamma_s(0) = s$, and that

$$\begin{aligned}\gamma'_s(t) &= d(L_s)_{\gamma(t)}(\gamma'(t)) \\ &= d(L_s)_{\gamma(t)}(X(\gamma(t))) \\ &= X(s \cdot \gamma(t)) \\ &= X(\gamma_s(t))\end{aligned}$$

Thus we need only show that γ is defined for all $t \in \mathbb{R}$. To do this, we first show that for all $t, s \in (-\epsilon, \epsilon)$ such that $t + s \in (-\epsilon, \epsilon)$

$$\gamma(s + t) = \gamma(s) \cdot \gamma(t) \tag{8}$$

Fix $s \in (-\epsilon, \epsilon)$. Let $\alpha(t) = \gamma(t + s)$. Then $\alpha(0) = \gamma(s)$ and by the chain rule, $\alpha'(t) = X(\alpha(t))$, so $\alpha(t)$ is an integral curve of X through $\gamma(s)$. On the other hand, let $\beta(t) = \gamma(s) \cdot \gamma(t)$. Then $\beta(0) = \gamma(s)$, and by a similar computation $\beta'(t) = X(\beta(t))$, thus $\beta(t)$ is also an integral curve of X through $\gamma(s)$. By uniqueness, $\alpha(t) = \beta(t)$. This proves (8).

Now suppose initially that $\gamma(t)$ is defined on $(-\epsilon, \epsilon)$. Define

$$\eta(t) = \gamma\left(\frac{\epsilon}{2}\right) \cdot \gamma\left(t - \frac{\epsilon}{2}\right)$$

for $t \in \left(-\frac{\epsilon}{2}, \frac{3\epsilon}{2}\right)$. By (8), $\eta(0) = \gamma(0)$ and

$$\begin{aligned}\eta'(t) &= dL_{\gamma(\frac{\epsilon}{2})}\left(\gamma'\left(t - \frac{\epsilon}{2}\right)\right) \\ &= dL_{\gamma(\frac{\epsilon}{2})}\left(X\left(\gamma\left(t - \frac{\epsilon}{2}\right)\right)\right) \\ &= X\left(\gamma\left(\frac{\epsilon}{2}\right) \cdot \gamma\left(t - \frac{\epsilon}{2}\right)\right) \\ &= X(\eta(t))\end{aligned}$$

So once again by the uniqueness of integral curves, $\gamma(t) = \eta(t)$ on $\left(-\frac{\epsilon}{2}, \epsilon\right)$. Thus if we set

$$\bar{\gamma}(t) = \begin{cases} \gamma(t) & t \in (-\epsilon, \epsilon) \\ \eta(t) & t \in \left(\frac{\epsilon}{2}, \frac{3\epsilon}{2}\right) \end{cases}$$

Then $\bar{\gamma}$ is an extension of γ past $t = \epsilon$. Observe that we could have just as easily extended γ past $t = -\epsilon$ by defining $\eta(t) = \gamma(-\frac{\epsilon}{2}) \cdot \gamma(t + \frac{\epsilon}{2})$. It is also clear that this process can be repeated indefinitely, so that γ can be extended to arbitrary large intervals about 0 and is therefore defined for all $t \in \mathbb{R}$. □

It is clear from this proof that the integral curve of a left invariant vector field X through the identity element is especially distinguished. In fact, (8) says that it is a (Lie group) homomorphism of the additive group of real numbers into G .

Definition 3.6. *Let G be a Lie group. A 1-parameter subgroup of G is a Lie group homomorphism $\gamma : \mathbb{R} \rightarrow G$ from the additive group of real numbers into G .*

Theorem 3.1. *Let G be a Lie group. For every tangent vector $v \in T_e G$, there is a unique 1-parameter subgroup of G , γ , such that $\gamma'(0) = v$.*

Proof. We have just shown existence, for if $v \in T_e G$ and V is the left invariant vector field obtained by extending v , then the integral curve of V through e is a 1-parameter subgroup. For uniqueness, we need the following fact: If G and H are Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, then every homomorphism φ from G into H induces a homomorphism of Lie algebras $d\varphi_e : \mathfrak{g} \rightarrow \mathfrak{h}$. If G is connected, and $\varphi, \psi : G \rightarrow H$ are homomorphisms of G into H such that their corresponding Lie algebra homomorphisms $d\varphi_e, d\psi_e$ are identical, then $\varphi = \psi$. (for a proof see [1])

Now suppose γ is a 1-parameter subgroup of G with $\gamma'(0) = v$, i.e. $d\gamma_0(\frac{\partial}{\partial t}(0)) = v$. Then for any $\lambda \in \mathbb{R}$, $d\gamma_0(\lambda \frac{\partial}{\partial t}(0)) = \lambda v$, so that the Lie algebra homomorphism $d\gamma_0$ is completely determined by v . Thus if η is another 1-parameter subgroup of G with $\eta'(0) = v$, then η induces the same Lie algebra homomorphism as γ , and since $\mathbb{R}$ is connected, $\gamma = \eta$. □

We are now in a position to introduce the exponential map. Given any $X \in \mathfrak{g}$, the integral curve of X through e is the unique 1-parameter subgroup of G having $X(e)$ as its tangent vector at e . It will be convenient to refer to this integral curve as $\exp_X$ from now on.

Definition 3.7. *The exponential map, $\exp : \mathfrak{g} \rightarrow G$ is the map given by $\exp(X) = \exp_X(1)$*

The exponential map provides the means to prove many powerful theorems linking a Lie group with its Lie algebra. However, an exhaustive treatment of Lie group theory is not our goal, so we choose rather to conclude this section by stating some important properties of the exponential map. (For more on the relationship of a Lie group and its Lie algebra see [1].)

Proposition 3.3. *Let G be a Lie group, $X \in \mathfrak{g}$, $s, t \in \mathbb{R}$ and $r \in G$.*

(i) $\exp(tX) = \exp_X(t)$ for all $t \in \mathbb{R}$

$$(ii) \exp((s+t)X) = \exp(sX) \cdot \exp(tX)$$

(iii) $L_r \circ \exp_X$ is the unique integral curve of X through r

(iv) Let $x_t: \mathbb{R} \times G \rightarrow G$ be the flow of X (since X is complete its flow is defined globally). Then $x_t = R_{\exp_X(t)}$

Proof. (i) We prove the more general fact $\exp_{tX}(s) = \exp_X(ts)$. Put $\alpha(s) = \exp_X(ts)$. Then by the chain rule, $\alpha'(s) = tX(\alpha(s))$, and $\alpha(0) = \exp_X(0) = e$. By uniqueness, $\alpha(s) = \exp_{tX}(s)$. Setting $s = 1$ yields the desired result.

(ii) By (i), (ii) simply says that $\exp_X$ is a homomorphism of $\mathbb{R}$ into G .

(iii) We have already proved (iii) in showing that every left invariant vector field is complete.

(iv) For any $s \in G$, x_t sends s to $\gamma_s(t)$, where γ_s is the integral curve of X through s . But by (iii), $\gamma_s(t) = s \cdot \exp_X(t)$. Thus $x_t(s) = R_{\exp_X(t)}(s)$. □

4 Riemannian Geometry

We now turn our attention to the Riemannian geometry of Lie groups. A natural question to ask is whether or not there is a connection between the exponential map just defined to the ordinary exponential map on Riemannian manifolds. From this point on we will be working towards the answer to this question.

Definition 4.1. Let G be a Lie group. A Riemannian metric $\langle \cdot, \cdot \rangle$ on G is said to be left invariant if every left translation is an isometry of $\langle \cdot, \cdot \rangle$. That is, if for each $s, t \in G$ and every $u, v \in T_t G$,

$$\langle u, v \rangle_t = \langle d(L_s)_t(u), d(L_s)_t(v) \rangle_{st}$$

Right invariant metrics are defined analogously.

Any inner product $\langle \cdot, \cdot \rangle_e$ on $T_e G$ can be extended to a left invariant metric on G by defining, for each $s \in G$ and all $u, v \in T_s G$,

$$\langle u, v \rangle_s = \langle d(L_{s^{-1}})_s(u), d(L_{s^{-1}})_s(v) \rangle_e \quad (9)$$

It is not difficult to check that this construction does define an actual Riemannian metric on G that is left invariant. Intuitively, the inner products vary smoothly across G because the left translations L_s "vary smoothly" with s (think of definition 3.1). Right invariant metrics can be constructed in a similar fashion. A metric that is both left and right invariant is said to be bi-invariant. An arbitrary Lie group always admits a left or right invariant metric via the above construction, however, it is not possible to ensure the existence of a bi-invariant metric without an additional assumption.

Theorem 4.1. *Every compact Lie group G admits a bi-invariant metric.*

We omit a proof of this result to avoid a discussion of differential forms. However, the basic idea of the construction is to use an "averaging" technique similar in spirit to (9). For more details see [2].

Remark. *If G is abelian, left and right translations coincide, so every left invariant metric is bi-invariant.*

As we shall see, the geometry of Lie groups with bi-invariant metrics is especially nice. This is mostly due to the following property of bi-invariant metrics:

Proposition 4.1. *Let G be a Lie group. If $\langle \cdot, \cdot \rangle$ is bi-invariant, then for every $X, Y, Z \in \mathfrak{g}$,*

$$\langle [X, Y], Z \rangle + \langle Y, [X, Z] \rangle = 0 \quad (10)$$

Proof. Let g be the metric tensor, i.e. $g(\cdot, \cdot) = \langle \cdot, \cdot \rangle$. Taking the Lie derivative of $g(Y, Z)$ with respect to X and using proposition 2.5, we have

$$\mathcal{L}_X(g(Y, Z)) = (\mathcal{L}_X g)(Y, Z) + g(\mathcal{L}_X Y, Z) + g(Y, \mathcal{L}_X Z) \quad (11)$$

By proposition 2.4, the left hand side is $X(g(Y, Z))$. The left invariance of g implies that the function $g(Y, Z)$ is the constant function taking the value $\langle Y(e), Z(e) \rangle_e$ everywhere, since for any $s \in G$,

$$\langle Y(s), Z(s) \rangle_s = \langle d(L_{s^{-1}})_s(Y(s)), d(L_{s^{-1}})_s(Z(s)) \rangle_e = \langle Y(e), Z(e) \rangle_e$$

Therefore

$$\mathcal{L}_X(g(Y, Z)) = 0$$

By proposition 3.3, the flow x_t of the left-invariant vector field X is right translation by $\exp_X(t)$. Since g is right invariant, all right translations are isometries, and thus X is a Killing field. By proposition 2.6, $(\mathcal{L}_X g)(Y, Z) = 0$. Therefore (11) becomes

$$g(\mathcal{L}_X Y, Z) + g(Y, \mathcal{L}_X Z) = 0$$

Using the fact that $\mathcal{L}_X Y = [X, Y]$, we see that this is exactly (10).

□

We note here that this proposition along with its converse (Proposition 4.2) characterizes bi-invariant metrics exactly as the left invariant metrics satisfying (10), and that the proofs of both of these propositions are original. Proposition 4.1 in its current state already allows us to describe the geometry of Lie groups with bi-invariant metrics in great detail. In what follows, G always refers to a Lie group with a bi-invariant metric and ∇ its Levi-Civita connection. Note that all of the following results will hold on compact Lie groups, as a consequence of theorem 4.1.

Theorem 4.2. *For every left invariant vector field X on G , $\nabla_X X = 0$*

Proof. Let $X, Y, Z \in \mathfrak{g}$ be arbitrary. Since ∇ is the Levi-Civita connection, it is symmetric and respects the metric, i.e. $\nabla_X Y - \nabla_Y X = [X, Y]$ and $Z\langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle$ for every $X, Y, Z \in \mathcal{X}(G)$. From proposition 4.1, we have

$$\begin{aligned} 0 &= \langle [X, Y], Z \rangle + \langle Y, [X, Z] \rangle \\ &= \langle \nabla_X Y - \nabla_Y X, Z \rangle + \langle Y, \nabla_X Z - \nabla_Z X \rangle \\ &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle - \langle \nabla_Y X, Z \rangle - \langle Y, \nabla_Z X \rangle \\ &= X\langle Y, Z \rangle - \langle \nabla_Y X, Z \rangle - \langle Y, \nabla_Z X \rangle \end{aligned}$$

The left invariance of $\langle \cdot, \cdot \rangle$ implies that $\langle Y, Z \rangle$ is a constant function, so $X\langle Y, Z \rangle = 0$ and

$$\langle \nabla_Y X, Z \rangle + \langle Y, \nabla_Z X \rangle = 0$$

Putting $X = Y$ and noting that $\langle X, \nabla_Z X \rangle = \frac{1}{2}Z\langle X, X \rangle = 0$, we get

$$\langle \nabla_X X, Z \rangle = 0$$

for every $Z \in \mathfrak{g}$. Hence $\nabla_X X = 0$. □

Corollary. *If $X, Y \in \mathfrak{g}$, then $\nabla_X Y = \frac{1}{2}[X, Y]$*

Proof. By Theorem 4.2, $\nabla_{X+Y}(X+Y) = 0$. Expanding this using linearity and then using the fact that ∇ is symmetric gives the desired relation. □

Let $X \in \mathfrak{g}$ and let $\gamma(t) = \exp_X(t)$. Then $\gamma'(t) = X(\gamma(t))$, so using theorem 4.2,

$$\frac{D\gamma'}{dt}(t) = \nabla_{\gamma'(t)}\gamma' = \nabla_{X(\gamma(t))}X = 0$$

Thus 1-parameter subgroups are geodesics. Indeed, since left translations are isometries, all left translates of 1-parameter subgroups are geodesics. It is simple to see that these must exhaust all the geodesics of G : Fix a point $s \in G$ and a vector $v \in T_s G$. The vector $d(L_{s^{-1}})_s(v)$ is in $T_e G$, and thus may be extended to a left invariant field X . Then $X(s) = v$, so $L_s \circ \exp_X(t)$ is the unique geodesic passing through s with tangent vector v .

We summarize the discussion and give an affirmative answer to the question posed at the beginning of section 4 in the following theorem

Theorem 4.3. *The geodesics of G coincide with the collection of left-translates of 1-parameter subgroups of G and the exponential map coincides with the usual exponential map of Riemannian geometry at the identity element.*

Corollary. *G is geodesically complete.*

As mentioned earlier, these results can be used to give an easy proof of the converse of proposition 4.1. (Note the only facts about the metric that were used to obtain the previous results were that it was left invariant and satisfied (10))

Proposition 4.2. *Let G be a Lie group. If a (left invariant) metric $\langle \cdot, \cdot \rangle$ on G satisfies*

$$\langle [X, Y], Z \rangle + \langle Y, [X, Z] \rangle = 0$$

for every $X, Y, Z \in \mathfrak{g}$, then it is bi-invariant.

Proof. Again taking the Lie derivative of $g(Y, Z)$ with respect to X , we have

$$\mathcal{L}_X(g(Y, Z)) = (\mathcal{L}_X g)(Y, Z) + g(\mathcal{L}_X Y, Z) + g(Y, \mathcal{L}_X Z)$$

But now by assumption, the last two terms on the right sum to zero, and the left invariance of g once again implies that the left hand side is zero. Hence

$$(\mathcal{L}_X g)(Y, Z) = 0$$

for every $Y, Z \in \mathfrak{g}$. By proposition 2.5, every $X \in \mathfrak{g}$ is a Killing field. It remains only to show that every right translation is the flow of some left invariant vector field. Fix $s \in G$. Since G is geodesically complete, there is a geodesic γ joining e to s . By uniqueness, this must be a 1-parameter subgroup. Let $X \in \mathfrak{g}$ be such that $\exp_X = \gamma$. Then $\exp_X(t) = s$ for some t , so $R_s = R_{\exp_X(t)}$ which is the flow of X . □

As a final application of these results, we give a bound on the sectional curvature of G . Recall that the Riemann curvature tensor $R(X, Y)Z$ is defined as

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

for $X, Y, Z \in \mathcal{X}(M)$. The sectional curvature K is the symmetric 2 - tensor

$$K(X, Y) = \langle R(X, Y)X, Y \rangle$$

The sectional curvature is a conceptual generalization of Gauss curvature of surfaces to higher dimensional manifolds. Roughly speaking, $K(X, Y)(p)$ measures how the plane in $T_p M$ spanned by $X(p)$ and $Y(p)$ changes as one moves an infinitesimal amount away from p in a fixed direction.

Theorem 4.4. *For every quadruple of left invariant vector fields X, Y, Z, W on G ,*

$$\langle R(X, Y)Z, W \rangle = \frac{1}{4} \langle [X, Y], [Z, W] \rangle \quad (12)$$

Proof. Using the fact that $\nabla_X Y = \frac{1}{2}[X, Y]$, we find that

$$\begin{aligned} R(X, Y)Z &= \frac{1}{4}[X, [Y, Z]] - \frac{1}{4}[Y, [X, Z]] - \frac{1}{2}[[X, Y], Z] \\ &= \frac{1}{4}[[X, Y], Z] \end{aligned}$$

Where the second equality is obtained by an application of the Jacobi identity. Rearranging the equality (10) yields $\langle [X, Y], Z \rangle = \langle X, [Y, Z] \rangle$, therefore

$$\begin{aligned} \langle R(X, Y)Z, W \rangle &= \frac{1}{4}\langle [[X, Y], Z], W \rangle \\ &= \frac{1}{4}\langle [X, Y], [Z, W] \rangle \end{aligned}$$

□

Corollary. *The sectional curvature of G , $K(X, Y) = \langle R(X, Y)X, Y \rangle$, is non-negative.*

Proof. This follows from theorem 4.4 since $\langle R(X, Y)X, Y \rangle = \langle [X, Y], [X, Y] \rangle \geq 0$ for all $X, Y \in \mathfrak{g}$.

□

At this point there are several possibilities. The reader may be interested in general Riemannian geometry, a full treatment of the general theory of Lie groups, or more on the Riemannian geometry of Lie groups. We recommend the texts by Warner [1] and Lee [4] for general Lie groups, and Do Carmo's text [3] for a reference to Riemannian geometry. Milnor's text [2] also contains most of the content included in section 4, plus more.

References

- [1] Frank W. Warner, *Foundations of Differentiable Manifolds and Lie groups*. Springer-Verlag, 1983.
- [2] John Milnor, *Morse Theory*. Princeton University Press, 1969.
- [3] Manfredo P. Do Carmo, *Riemannian Geometry*. Birkhäuser, 1992.
- [4] John M. Lee, *Introduction to Smooth Manifolds*. Springer-Verlag, 2003.